JMIR Cancer
Volume 10 (2024)

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

Contents
Reviews
Digital Health Psychosocial Intervention in Adult Patients With Cancer and Their Families: Systematic
Review and Meta-Analysis (e46116)
Yingzi Zhang, Marie Flannery, Zhihong Zhang, Meghan Underhill-Blazey, Melanie Bobry, Natalie Leblanc, Darcey Rodriguez, Chen Zhang. . .
4

Machine Learning Approaches to Predict Symptoms in People With Cancer: Systematic Review (e52322)
Nahid Zeinali, Nayung Youn, Alaa Albashayreh, Weiguo Fan, Stéphanie Gilbertson White. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

24

Implementation of Health IT for Cancer Screening in US Primary Care: Scoping Review (e49002)
Constance Owens-Jasey, Jinying Chen, Ran Xu, Heather Angier, Amy Huebschmann, Mayuko Ito Fukunaga, Krisda Chaiyachati, Katharine
Rendle, Kim Robien, Lisa DiMartino, Daniel Amante, Jamie Faro, Maura Kepper, Alex Ramsey, Eric Bressman, Rachel Gold. . . . . . . . . . . . . . . . . .

274

Viewpoints
Need for Culturally Competent and Responsive Cancer Education for African Immigrant Families and Youth
Living in the United States (e53956)
Olufunmilola Abraham, Adeola Agoke, Kazeem Sanuth, Abimbola Fapohunda, Motolani Ogunsanya, Megan Piper, Amy Trentham-Dietz. . . .
4
1

An Introduction to the OutSMART Cancer Serious Game: Current and Future Directions (e56168)
Olufunmilola Abraham, Tyler McCarthy. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

53

“Notification! You May Have Cancer.” Could Smartphones and Wearables Help Detect Cancer Early?
(e52577)
Suzanne Scott, Matthew Thompson. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

379

Original Papers
Iterative Patient Testing of a Stimuli-Responsive Swallowing Activity Sensor to Promote Extended User
Engagement During the First Year After Radiation: Multiphase Remote and In-Person Observational Cohort
Study (e47359)
Eileen Shinn, Adam Garden, Susan Peterson, Dylan Leupi, Minxing Chen, Rachel Blau, Laura Becerra, Tarek Rafeedi, Julian Ramirez, Daniel
Rodriquez, Finley VanFossen, Sydney Zehner, Patrick Mercier, Joseph Wang, Kate Hutcheson, Ehab Hanna, Darren Lipomi. . . . . . . . . . . . . . . . . . .

58

JMIR Cancer 2024 | vol. 10 | p.1

XSL• FO
RenderX


<table>
<thead>
<tr>
<th>Title</th>
<th>E-Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Design, Development, and Evaluation of a Mobile Solution to Improve Medication Adherence in Cancer: Design Science Research Approach (e46979)</td>
<td>75</td>
</tr>
<tr>
<td>Telehealth With Comprehensive Live-Fed Real-World Data as a Patient Care Platform for Lung Cancer: Implementation and Evaluation Study (e45331)</td>
<td>96</td>
</tr>
<tr>
<td>Applying the Unified Theory of Acceptance and Use of Technology to Identify Factors Associated With Intention to Use Teledelivered Supportive Care Among Recently Diagnosed Breast Cancer Survivors During COVID-19 in Hong Kong: Cross-Sectional Survey (e51072)</td>
<td>110</td>
</tr>
<tr>
<td>Usability and Preliminary Efficacy of an Adaptive Supportive Care System for Patients With Cancer: Pilot Randomized Controlled Trial (e49703)</td>
<td>129</td>
</tr>
<tr>
<td>Toxic Relationships Described by People With Breast Cancer on Reddit: Topic Modeling Study (e48860)</td>
<td>149</td>
</tr>
<tr>
<td>Exploring Web-Based Information and Resources That Support Adolescents and Young Adults With Cancer to Resume Study and Work: Environmental Scan Study (e47944)</td>
<td>162</td>
</tr>
<tr>
<td>Evaluation of the e–Mental Health Intervention Make It Training From Patients’ Perspectives: Qualitative Analysis Within the Reduct Trial (e53117)</td>
<td>175</td>
</tr>
<tr>
<td>Differing Content and Language Based on Poster-Patient Relationships on the Chinese Social Media Platform Weibo: Text Classification, Sentiment Analysis, and Topic Modeling of Posts on Breast Cancer (e51332)</td>
<td>190</td>
</tr>
<tr>
<td>Evaluating the Quality of Cancer-Related WeChat Public Accounts: Cross-Sectional Study (e52156)</td>
<td>205</td>
</tr>
<tr>
<td>Engagement With a Relaxation and Mindfulness Mobile App Among People With Cancer: Exploratory Analysis of Use Data and Self-Reports From a Randomized Controlled Trial (e52386)</td>
<td>213</td>
</tr>
<tr>
<td>Heart Rate Monitoring Among Breast Cancer Survivors: Quantitative Study of Device Agreement in a Community-Based Exercise Program (e51210)</td>
<td>232</td>
</tr>
<tr>
<td>Do Measures of Real-World Physical Behavior Provide Insights Into the Well-Being and Physical Function of Cancer Survivors? Cross-Sectional Analysis (e53180)</td>
<td>242</td>
</tr>
</tbody>
</table>
Improving Concordance Between Clinicians With Australian Guidelines for Bowel Cancer Prevention Using a Digital Application: Randomized Controlled Crossover Study (e46625)
Tsai-Wing Ow, Olga Sukacheva, Peter Bampton, Guruparan lyngkaran, Christopher Rayner, Edmund Tse ................................................................. 260

Pediatric Cancer Communication on Twitter: Natural Language Processing and Qualitative Content Analysis (e52061)
Nancy Lau, Xin Zhao, Alison O'Daffer, Hannah Weissman, Krysta Barton ....................................................................................................... 291

Digital Smoking Cessation Intervention for Cancer Survivors: Analysis of Predictors and Moderators of Engagement and Outcome Alongside a Randomized Controlled Trial (e46303)
Rosa Andree, Aija Mujicic, Wouter den Hollander, Margriet van Laar, Brigitte Boon, Rutger Engels, Matthijs Blankers ........................................ 305

Assessing the Quality, Privacy, and Security of Breast Cancer Apps for Arabic Speakers: Systematic Search and Review of Smartphone Apps (e48428)
Dari Alhuwail, Aisha Alhouti, Latifah Alsarhan ............................................................................................................................... 316

Comparison of the Real-World Reporting of Symptoms and Well-Being for the HER2-Directed Trastuzumab Biosimilar Ogivri With Registry Data for Herceptin in the Treatment of Breast Cancer: Prospective Observational Study (OGIPRO) of Electronic Patient-Reported Outcomes (e54178)
Andreas Trojan, Sven Roth, Ziad Atassi, Michael Kiessling, Reinhard Zenhaeusern, Yannick Kadavany, Johannes Schumacher, Gerd Kullak-Ublick, Matti Aspro, Alexandru Eniu ................................................................. 328

Variation in Trust in Cancer Information Sources by Perceptions of Social Media Health Mis- and Disinformation and by Race and Ethnicity Among Adults in the United States: Cross-Sectional Study (e54162)
Jim Stimpson, Sungchul Park, Sandi Pruitt, Alexander Ortega ........................................................................................................ 339

Yue Zheng, Ailin Zhao, Yuqi Yang, Laduona Wang, Yifei Hu, Ren Luo, Yijun Wu .............................................................................................. 349

Patterns of Prescription Medication Use Before Diagnosis of Early Age-Onset Colorectal Cancer: Population-Based Descriptive Study (e50402)
Vienna Cheng, Eric Sayre, Vicki Cheng, Ria Gang, Sharlene Gill, Ameer Farooq, Mary De Vera ........................................................................ 369

Research Letter

Web-Based Scaffolds: The Feasibility of a Constructivist Approach to Oncology Fellow Learning (e52501)
Sam Brondfield, Matthew Schwede, Tyler Johnson, Shagun Arora ........................................................................................................ 387
Review

Digital Health Psychosocial Intervention in Adult Patients With Cancer and Their Families: Systematic Review and Meta-Analysis

Yingzi Zhang1*, PhD; Marie Flannery2, PhD; Zhihong Zhang2, MSc; Meghan Underhill-Blazey2, PhD; Melanie Bobry2, MSc; Natalie Leblanc2, MPH, PhD; Darcey Rodriguez3, MLiS; Chen Zhang2*, MPH, PhD

1Magnet Program and Nursing Research Department, UT Southwestern Medical Center, Dallas, TX, United States
2School of Nursing, University of Rochester Medical Center, Rochester, NY, United States
3Edward G Miner Library, University of Rochester Medical Center, Rochester, NY, United States
*these authors contributed equally

Abstract

Background: Patients with cancer and their families often experience significant distress and deterioration in their quality of life. Psychosocial interventions were used to address patients’ and families’ psychosocial needs. Digital technology is increasingly being used to deliver psychosocial interventions to patients with cancer and their families.

Objective: A systematic review and meta-analysis were conducted to review the characteristics and effectiveness of digital health interventions on psychosocial outcomes in adult patients with cancer and their family members.

Methods: Databases (PubMed, Cochrane Library, Web of Science, Embase, CINAHL, PsycINFO, ProQuest Dissertations and Theses Global, and ClinicalTrials.gov) were searched for randomized controlled trials (RCTs) or quasi-experimental studies that tested the effects of a digital intervention on psychosocial outcomes. The Joanna Briggs Institute’s critical appraisal checklists for RCTs and quasi-experimental studies were used to assess quality. Standardized mean differences (ie, Hedges $g$) were calculated to compare intervention effectiveness. Subgroup analysis was planned to examine the effect of delivery mode, duration of the intervention, type of control, and dosage on outcomes using a random-effects modeling approach.

Results: A total of 65 studies involving 10,361 patients (mean 159, SD 166; range 9-803 patients per study) and 1045 caregivers or partners (mean 16, SD 54; range 9-244 caregivers or partners per study) were included in the systematic review. Of these, 32 studies were included in a meta-analysis of the effects of digital health interventions on quality of life, anxiety, depression, distress, and self-efficacy. Overall, the RCT studies’ general quality was mixed (applicable scores: mean 0.61, SD 0.12; range 0.38-0.91). Quasi-experimental studies were generally of moderate to high quality (applicable scores: mean 0.75, SD 0.08; range 0.63-0.89). Psychoeducation and cognitive-behavioral strategies were commonly used. More than half (n=38, 59%) did not identify a conceptual or theoretical framework. Most interventions were delivered through the internet (n=40, 62%). The median number of intervention sessions was 6 (range 1-56). The frequency of the intervention was highly variable, with self-paced (n=26, 40%) being the most common. The median duration was 8 weeks. The meta-analysis results showed that digital psychosocial interventions were effective in improving patients’ quality of life with a small effect size (Hedges $g$=0.05, 95% CI –0.01 to 0.10; $I^2$=42.7%; $P$=.01). The interventions effectively reduced anxiety and depression symptoms in patients, as shown by moderate effect sizes on Hospital Anxiety and Depression Scale total scores (Hedges $g$=–0.72, 95% CI –1.89 to 0.46; $I^2$=97.6%; $P<.001$).

Conclusions: This study demonstrated the effectiveness of digital health interventions on quality of life, anxiety, and depression in patients. Future research with a clear description of the methodology to enhance the ability to perform meta-analysis is needed. Moreover, this study provides preliminary evidence to support the integration of existing digital health psychosocial interventions in clinical practice.

Trial Registration: PROSPERO CRD42020189698; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=189698
Introduction

Cancer is often associated with psychological distress in patients and their family members. Emerging evidence shows that psychological distress contributes to cancer mortality [1,2]. Given that over 2 million new cancer cases are expected to be diagnosed in 2024 in the United States, psychosocial distress is a significant public health problem [3]. Psychosocial distress can be triggered by many challenges, such as decision-making regarding treatment, self-care challenges due to side effects from cancer treatment, maintaining work-life balance, and financial burden. A large body of research documents the negative influence of a cancer diagnosis and treatment on a patient’s experience, including depression, anxiety, and decreased quality of life [4,5]. Cancer not only affects the patient but also imposes changes on the family [6]. Family members, who often assume caregiving roles to complement the roles of the health care team, often experience deteriorating quality of life and significant psychological distress [7,8]. For many years, researchers have examined psychosocial interventions addressing patients' and family members' needs to help maintain psychosocial well-being and quality of life during the cancer experience [9-12].

Increasingly, studies have used digital technology to deliver psychosocial interventions. In this report, we refer to digital health intervention as the use of digital, mobile, and wireless technologies to deliver an intervention. Digital health interventions have gained popularity due to their geographic accessibility, self-paced nature, user-friendly design, up-to-date information provision, and time-sensitive interaction with health care providers [13,14]. Further, digital interventions have significant potential for reaching people, mainly in rural areas or people with limited mobility [15]. There are various delivery modes for digital interventions, such as smartphone apps, websites, the internet, and virtual reality. There are also drawbacks, including concerns related to security and privacy and inaccessibility for people without smart device ownership. Psychosocial interventions may incorporate various components, such as communication skills training, cognitive behavioral therapy, patient education, peer support, and problem-solving training [16].

Despite the plethora of individual research studies, a synthesis of digital psychosocial interventions for patients with cancer and their families is needed to provide a summary of existing evidence regarding the effects of interventions and provide directions for future research and clinical practice. A range of systematic reviews have examined digital health psychosocial interventions for patients with cancer [17-22] and their family members [23,24]. However, these reviews have limitations. For example, some reviews primarily focused on a specific population, such as individuals with breast [17] or prostate cancer [19,20]; a particular delivery mode, such as internet-based [17,23,24]; or a specific psychosocial outcome, such as quality of life or psychological distress [21,22]. In addition, Slev et al [25] synthesized evidence from systematic reviews of interventions delivered through computers or the internet for patients with cancer and their caregivers; however, the authors failed to quantify the effectiveness of interventions across studies using advanced statistical techniques, such as a meta-analysis. To date, no studies have used meta-analytical strategies to quantify the impact of digital health interventions on psychosocial outcomes in patients with cancer and family members. To fill these gaps, we conducted a systematic review and meta-analysis to comprehensively review the characteristics and effectiveness of digital psychosocial interventions on psychosocial outcomes across different available delivery modes in adult patients with cancer and their family members.

The specific aims were to answer the following questions:

1. What are the characteristics of digital psychosocial interventions for adult patients and families living with cancer? (ie, intervention component, theoretical or conceptual framework, tailored or standardized, mode of delivery, prescribed dosage, duration of the intervention, and actual dosage)?
2. What is the efficacy of interventions on psychosocial outcomes for adult individuals diagnosed with cancer and their family members and associated factors (ie, delivery mode, control condition, and dosage, including the number of sessions, frequency, and duration)?

Methods

The review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist [26].

Study Identification

The medical librarian (DR) and first author (YZ) worked together to identify search terms to build a comprehensive search strategy (Multimedia Appendix 1). Using controlled vocabulary and keywords when available, the search strategy was executed in the following databases: PubMed, Cochrane Library, Web of Science, Embase, CINAHL, PsycINFO, ProQuest Dissertations and Theses Global, and ClinicalTrials.gov. The results were limited to the English language and those published from each resource’s inception until March 2019, when the search was completed. An initial limited search of PubMed and CINAHL was undertaken, followed by an analysis of the text words in the abstract and the index terms used to describe the article. Relevancy was determined by the first author (YZ) and medical librarian (DR). A second search was undertaken across all included databases using all identified keywords and index terms.
Study Selection
The inclusion criteria were studies that (1) included adult patients (≥18 years of age with any cancer diagnosis) or their adult family members (eg, partner, caregiver, adult children, parent, or relative); (2) tested a digital health psychosocial intervention, which was defined as any nonpharmacological therapeutic intervention that addressed the psychological, social, personal, or relational adjustment needs associated with cancer through a digital health mechanism (eg, application and website); (3) measured at least 1 psychosocial outcome; and (4) used an experimental (randomized controlled trial [RCT]) or a quasi-experimental design. Studies were excluded if they enrolled pediatric patients with cancer; were review articles, letters to the editor, editorial reports, case reports, or commentaries; were published as abstracts only; and were not published in English. For meta-analysis, we excluded articles that did not provide data or when only a single study included the outcome measure.

After removing duplicates, the first author (YZ) read all titles and abstracts to identify articles based on inclusion and exclusion criteria. The full texts of all included articles were then screened independently by 2 reviewers (master’s-level or above), and final decisions were made based on consensus. Finally, articles identified in the search were imported to Endnote X8 (Clarivate Analytics).

Data Extraction and Management
A Microsoft Excel (Microsoft Corporation) spreadsheet was used to record information [27], including the description of the interventions (eg, theory basis, mode of delivery, content, actual dosage, planned dosage, standardized, or tailored), study sample (eg, age, sex, education, race, ethnicity, and cancer diagnosis), study characteristics (eg, design, randomization method, and control condition), intervention outcome variables and measurements, follow-ups, and quantitative data (ie, mean, SD, and sample size). Dosage was described as the number of intervention sessions, frequency, and duration of access to intervention. A standardized intervention was defined as all participants receiving the same intervention, while a tailored intervention involved customization of the intervention based on individual characteristics or needs [28]. We defined the prescribed dosage as the intended treatment dose, including the number of intervention sessions, frequency, and total length according to the study protocol. A codebook was created for data extraction, and the team’s decisions were tracked and recorded. All authors extracted data from 3 articles to pilot-test the spreadsheet. The research team discussed any ambiguity, resolved differences in interpretation, and modified the data extraction spreadsheet. Subsequently, each article underwent independent data extraction by YZ and another author (6 trained reviewers). The research team met throughout the study period every other week to resolve any discrepancies. A total of 15 original study authors were contacted to request missing information (eg, mean, SD, and sample size), and no additional data were received.

Assessment of Methodological Quality
The reviewers assessed the included studies for methodological rigor using standardized critical appraisal instruments from the 13-item Joanna Briggs Institute (JBI) Critical Appraisal Checklist for RCT and the 9-item JBI Critical Appraisal Checklist for quasi-experimental studies [29]. Reviewers answered each risk of bias item as “yes” (score=1), “no” (score=0), “unclear” (score=0), or “not applicable.” Possible composite scores ranged from 0 to 9 for quasi-experimental studies and 0-13 for RCTs, with higher scores indicating less risk of bias and better study quality. The applicable score (range 0-1) was calculated by dividing the composite score by the maximum score possible after subtracting any “not applicable” responses [30]. All studies were double-coded, and any disagreements were resolved through discussion with the research team [26].

Data Synthesis and Meta-Analysis

Data Synthesis
Data synthesis was completed on all articles that met the inclusion criteria. Only primary study results were included if multiple articles were published from the same intervention study. Simple descriptive statistics (ie, mean, SD, frequency, and percentage) were used to summarize study characteristics (eg, study design and participant characteristics) and key features of interventions (ie, theory, mode of delivery, number of sessions, frequency, and total length). Intervention content was grouped and narratively summarized according to the description of the intervention components.

Meta-Analytical Procedure
An a priori decision was made to only include studies in the meta-analysis if at least 2 studies used the same instrument to assess the same psychosocial outcome [31]. Standardized mean differences (ie, Hedges g) were calculated to compare intervention effectiveness across studies that used different scales or measurements. Mean differences between the scores before the intervention and the follow-up assessment after the intervention were calculated for pre-post interventions. Similarly, for the RCT studies, the results from follow-up in each study were selected and analyzed using difference scores from before and after the intervention for both intervention and control groups, with the pooled SDs. We computed the overall effect size across different time points for studies with multiple follow-ups. By doing so, we captured the time-varying effect on intervention effectiveness [31]. The overall effect (including all information across all time points) and time-varying effects, including the interim effect (during the intervention period), immediate effect (after the intervention), short-term effect (follow-up ≤8 weeks after completion of the intervention), and long-term effect (follow-up >8 weeks after completion of the intervention), were calculated. A cutoff of 8 weeks was chosen because it was the median length of the follow-up period across the included studies.

To assess study heterogeneity, the $I^2$ statistic was examined. The $I^2$ statistic quantifies the proportion of total variance across studies caused by a fundamental difference between trials rather than chance. An $I^2$ statistic of <25% indicates low heterogeneity,
between 25% and 75% indicates moderate heterogeneity, and >75% indicates high heterogeneity [32]. Lower heterogeneity is better. Funnel plots (ie, to visually assess the asymmetry) and Egger test (ie, to test the asymmetry statistically) assessed publication bias [33]. In funnel plots, if points are distributed equally between positive and negative effects, bias is lacking; variability is expected to be greater near the bottom of the chart among smaller sample size studies. For the analysis of data from studies with more than 1 digital psychosocial intervention group, we compared each digital psychosocial intervention group to the control group separately. Additionally, subgroup analysis was planned based on the review’s focus on examining the effect of delivery mode, type of control condition, and dosage on outcomes. Furthermore, we performed sensitivity analyses by including and excluding studies with extreme weights in the analyses. We used the DerSimonial-Laird random-effects model to weight and pool the individual estimates to capture variance across different studies, as all included studies were conducted in heterogeneous populations across various settings [34]. We performed all statistical and meta-analyses using STATA (version 17; StataCorp LLC).

### Results

#### Search Results

After removing duplicates, a total of 2108 studies were identified. Figure 1 shows a flow diagram of studies identified, screened, included, and excluded from this systematic review and meta-analysis. After screening titles and abstracts and applying inclusion and exclusion criteria, a total of 70 records with 65 unique studies (for multiple manuscripts published from the same intervention study, only primary manuscripts were included) were included in the systematic review [35-99] and 32 studies [35,36,38,40,41,43,44,47,50,53-55,57,64,66,68-72, 74,77,80,82,83,86,88,91,93,95,96,99] with available data were included in the meta-analysis. A total of 33 studies were excluded from the meta-analysis because either data were unavailable to calculate the effect size (n=14) [42,46,51,62,73,76,78,79,81,84,87,89,94,98] or no other study used the same measure (n=19) [37,39,45,48,49,52, 56,58-61,63,65,67,75,85,90,92,97].

![Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.](https://cancer.jmir.org/2024/1/e46116)

### Study Characteristics

#### Overview

Of the 65 studies, 48 (74%) were RCTs [35-37, 39-50,52-57,59-65,72,73,76-82,85,89,91,93-99], and 17 (26%) were quasi-experimental [38,51,58,66-71,74,75,83,84,86-88,92]. More than half (n=37, 58%) of the studies were conducted in the United States [36-40,44,48-50,54,56-61,63, 65,67-69,71,72,74,77,79,80,83-85,87-90,94,97,98], and the rest were from the Netherlands (n=9, 14%) [35,41,43,47,64,70,91,92,96], Australia (n=5, 8%) [45,66,75,86,95], and other countries (eg, Denmark and Ireland). In total, 10,361 patients (mean 159, SD 166; range 9-803
patients per study) were included: 7098 female patients and 3263 male patients; 1045 caregivers or partners were enrolled (mean 16, SD 54; range 9-244 caregivers or partners per study), including 781 female individuals and 264 male individuals. The average age of patients ranged from 39.9 to 72 years, and the average age of caregivers or partners ranged from 51.5 to 58.8 years. In the 33 studies that provided information about race and ethnicity, most patients (n=3495, 90%) and family members (n=259, 97%) were described as “White” or “Caucasian.” The cancer diagnoses varied across studies, with the most prevalent being breast cancer (n=24, 37%) [35,37,38,42,44,52,54,58,60,61,64-66,70-72,74,76,80,82,84,91,93,95], mixed cancer diagnosis (n=19, 29%) [36,40,45,47,50,51,53,57,62,69,75,77,78,81,83,94,96,97,99], and prostate cancer (n=7, 11%) [39,48,56,67,85,88,99]. The attrition rate ranged from 0% to 76%, with a median of an attrition rate of 4.4% to 94.2%, with a median of a 59.5% (mean 56%, SD 13.7%). The recruitment rate ranged from 0% to 76%, with a median of a recruitment rate of an attrition rate of 4.4% to 94.2%, with a median of a recruitment rate of 59.5% (mean 56%, SD 13.7%).

Aim 1: Intervention Characteristics

Overview

There was large heterogeneity in intervention components, theoretical or conceptual framework, type of intervention (ie, tailored or standardized), mode of delivery, prescribed dosage (ie, number of sessions, frequency, and length), and received dosage (Table S2 in Multimedia Appendix 2).

Intervention Components

A total of 37 (57%) out of 65 studies included a single intervention component [36,38,40,44,45,47,51,52,55-57,63-66,68,69,71-77,79,80,82-84,86,87,90,91,93,94,98], 13 (20%) studies included 2 intervention components [35,41,43,46,48,50,62,67,78,85,88,95,96], and 15 (23%) studies included 3-5 intervention components [37,42,49,53,54,58-61,70,81,89,92,97,99]. The most common intervention components were information and resources, or psychoeducation (n=29, 45%) [35,37,39-43,46,48,49,52,56-58,61,70,81,82,87-90,92,95,97], and cognitive-behavioral strategies (n=20, 31%) [44,45,47,50,54,57,63,64,67,68,71,74,75,80,85-87,89,91,98].

Theoretical or Conceptual Framework

More than half (n=38, 59%) of the included studies did not identify a conceptual or theoretical framework [37,41,42,45,47,51,52,54,55,57,63,66,69,71,73,75,76,78,81,85,90,92,93,95,97,98].

Standardized or Tailored Intervention

Of the 65 studies, 26 (40%) included both standardized and tailored interventions [37,39,42,43,46,47,49,53,55,59-61,64,68,70,73-75,77-79,81,85,89,92,94], 28 (43%) studies included only standardized interventions [36,38,41,44,50,51,54,56-58,63,65-67,69,71,72,76,80,82-84,87,88,91,93,95,98], and 11 (17%) studies had only tailored interventions [35,40,45,48,52,62,86,90,96,97,99].

Modes of Delivery

The majority of studies conducted interventions through an internet website (n=40, 62%) [35-37,39-50,53,55,59-62,64,67,68,70,72,75,77,81,82,85,88-91,93,95-99] or smartphone app (n=8) [43,52,56,57,63,69,80,87]. A total of 7 (11%) studies conducted interventions through virtual reality [51,66,73,76,78,83,84], 3 (5%) studies conducted interventions through telehealth [54,74,79], and 2 (3%) studies through a computer program [38,65]. Electronic health information systems [92], interaction portals [58], and videoconferences [86] were each used in 1 study. Overall, 2 studies used multimodal interventions delivered through the combination of either telephone and videoconference [94] or internet and telephone [61].

Dosage

The dosage prescribed and received were highly variable. The number of intervention sessions ranged from 1 to 56, with a median of 6. A total of 27 (42%) studies did not specify the prescribed dosage; 19 (29%) only stated the number of days participants had access to the intervention [35,36,42,49,56,59-62,67,70,80,81,85,87,89,92,95,99] and 8 (12%) did not provide information on the prescribed dosage [37,39,40,52,55,58,72,73]. Frequency was highly variable, with
self-paced (n=26, 40%) as the most common [35, 36, 42, 45, 49, 50, 56, 59-63, 67, 70, 80, 81, 85, 87-90, 92, 95-97, 99], meaning no specific intervention frequency was defined and the intervention content was available throughout the study period. The other common frequencies of intervention sessions were weekly (n=17, 26%) [38, 41, 44, 45, 47, 53, 54, 64, 68, 71, 74, 75, 77, 86, 91, 94, 98] and 1-time intervention sessions (n=8, 12%) [48, 65, 66, 76, 78, 82-84]. The median length of the intervention was 8 weeks, with the length ranging from 1 hour (ie, use of the intervention on an iPad for an hour) to 24 months.

Received dosage was defined as the uptake of the intervention by the participants. A total of 18 (28%) studies did not report the received dosage [36, 37, 39, 48, 52, 56, 58, 65, 69, 75, 76, 78, 79, 82, 85, 86, 93, 94]. Various information was reported, including attendance rate, number of times participants used the app, frequency with which participants logged into the website, number of website pages reviewed, skill practice time, and intervention session completion rate. Most of the interventions (n=43, 66%) were self-delivered without an interventionist, with self-paced being most common [35-37, 39, 40, 42, 44, 45, 48-53, 55, 57-60, 62, 63, 67, 69-73, 75, 76, 78, 80-84, 87, 88, 91-93, 95, 96, 99].

Aim 2: Effects on Patients’ and Family Members’ Psychosocial Outcomes

Patients’ Outcomes

Overview

A meta-analysis was conducted on 32 studies. Overall, 5 outcomes were examined. A summary of the interventions’ overall effect sizes; time-varying effect sizes for quality of life, anxiety, depression, distress, and self-efficacy; and heterogeneity statistics for each outcome is displayed in Table 1. The forest plots for overall effect sizes and time-varying effects are displayed in Multimedia Appendix 3. The funnel plots for overall effect sizes and time-varying effects are displayed in Multimedia Appendix 4.
<table>
<thead>
<tr>
<th>Population, outcome, measure, and value</th>
<th>Effect at different time points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Patient</td>
<td></td>
</tr>
<tr>
<td>QOL&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>FACT-B&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Pool ES&lt;sup&gt;d&lt;/sup&gt;, Hedges g (95% CI)</td>
<td>0.13 (–0.05 to 0.31)</td>
</tr>
<tr>
<td>$I^2$</td>
<td>62.3</td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2$ ($df$)</td>
<td>10.61 (4)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.03</td>
</tr>
<tr>
<td>FACT-G&lt;sup&gt;f&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Pool ES, Hedges g (95% CI)</td>
<td>–0.04 (–0.17 to 0.09)</td>
</tr>
<tr>
<td>$I^2$</td>
<td>0</td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2$ ($df$)</td>
<td>1.91 (4)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.43</td>
</tr>
<tr>
<td>QLQ-30&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Pool ES, Hedges g (95% CI)</td>
<td>0.05 (–0.04 to 0.14)</td>
</tr>
<tr>
<td>$I^2$</td>
<td>58.4</td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2$ ($df$)</td>
<td>19.95 (6)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.03</td>
</tr>
<tr>
<td>SF36&lt;sup&gt;h&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Pool ES, Hedges g (95% CI)</td>
<td>0.03 (–0.10 to 0.15)</td>
</tr>
<tr>
<td>$I^2$</td>
<td>14.4</td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2$ ($df$)</td>
<td>8.41 (8)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.31</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>Pool ES, Hedges g (95% CI)</td>
<td>0.05 (–0.01 to 0.10)</td>
</tr>
<tr>
<td>$I^2$</td>
<td>42.7</td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2$ ($df$)</td>
<td>48.12 (20)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.01</td>
</tr>
<tr>
<td>Anxiety and depression</td>
<td></td>
</tr>
<tr>
<td>HADS&lt;sup&gt;i&lt;/sup&gt; total score</td>
<td></td>
</tr>
<tr>
<td>Pool ES, Hedges g (95% CI)</td>
<td>–0.72 (–1.89 to 0.46)</td>
</tr>
<tr>
<td>$I^2$</td>
<td>97.6</td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2$ ($df$)</td>
<td>165.82 (14)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

<sup>a</sup> Pooled ES, Hedges g (95% CI)

<sup>b</sup> FACT = Functional Assessment of Cancer Therapy, QOL = quality of life

<sup>c</sup> FACT-B = Functional Assessment of Cancer Therapy-Breast

<sup>d</sup> Hedges g = standardized mean difference

<sup>e</sup> Heterogeneity test

<sup>f</sup> FACT-G = Functional Assessment of Cancer Therapy-General

<sup>g</sup> QLQ-30 = Quality of Life Questionnaire

<sup>h</sup> SF36 = Medical Outcomes Study 36-item Short-Form Health Survey

<sup>i</sup> HADS = Hospital Anxiety and Depression Scale

<sup>j</sup> Total score

Note: The table provides a summary of the meta-analysis results, including effect sizes, confidence intervals, and heterogeneity statistics for different outcomes and time points.
<table>
<thead>
<tr>
<th>Population, outcome, measure, and value</th>
<th>Effect at different time points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall(^a)</td>
</tr>
<tr>
<td><strong>HADS-depression</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>-0.13 (--0.23 to -0.02)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>0</td>
</tr>
<tr>
<td>Heterogeneity, (\chi^2 (df))</td>
<td>4.17 (7)</td>
</tr>
<tr>
<td>(P) value</td>
<td>.73</td>
</tr>
<tr>
<td><strong>CESD(^j)</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>0.10 (--0.10 to 0.30)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>0</td>
</tr>
<tr>
<td>Heterogeneity, (\chi^2 (df))</td>
<td>0.99 (4)</td>
</tr>
<tr>
<td>(P) value</td>
<td>.91</td>
</tr>
<tr>
<td><strong>PHQ9(^k)</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>-0.05 (--0.17 to 0.08)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>0</td>
</tr>
<tr>
<td>Heterogeneity, (\chi^2 (df))</td>
<td>0.78 (1)</td>
</tr>
<tr>
<td>(P) value</td>
<td>.38</td>
</tr>
<tr>
<td><strong>Multiple scales</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>0.32 (--0.35 to 0.99)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>95</td>
</tr>
<tr>
<td>Heterogeneity, (\chi^2 (df))</td>
<td>19.86 (1)</td>
</tr>
<tr>
<td>(P) value</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>0.03 (--0.10 to 0.16)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>60.9</td>
</tr>
<tr>
<td>Heterogeneity, (\chi^2 (df))</td>
<td>40.77 (16)</td>
</tr>
<tr>
<td>(P) value</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Anxiety</strong></td>
<td></td>
</tr>
<tr>
<td><strong>HADS-anxiety</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>0.32 (--0.20 to 0.84)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>94.3</td>
</tr>
<tr>
<td>Heterogeneity, (\chi^2 (df))</td>
<td>123.33 (7)</td>
</tr>
<tr>
<td>(P) value</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>SATI(^l)</strong></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>-0.19 (--0.41 to 0.04)</td>
</tr>
<tr>
<td>(I^2)</td>
<td>26.8</td>
</tr>
</tbody>
</table>

\(^a\) Pooled ES, Hedges g (95% CI)

\(^j\) CESD

\(^k\) PHQ9

\(^l\) SATI
<table>
<thead>
<tr>
<th>Population, outcome, measure, and value</th>
<th>Effect at different time points</th>
<th>Overalla</th>
<th>Immediate</th>
<th>Interim</th>
<th>Short</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneity, $\chi^2 (df)$</td>
<td>5.46 (4)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.24</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>0.12 ( –0.19 to 0.43)</td>
<td>–0.10 (–0.19 to 0)</td>
<td>–0.04 (–0.19 to 0.12)</td>
<td>–0.13 (–0.43 to 0.17)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$I^2$</td>
<td>90.2</td>
<td>6.7</td>
<td>35.1</td>
<td>10.5</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2 (df)$</td>
<td>132.99 (13)</td>
<td>13.94 (13)</td>
<td>6.16 (4)</td>
<td>1.12 (1)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>.38</td>
<td>.19</td>
<td>.29</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Distress</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DTm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>0.98 ( –0.18 to 2.14)</td>
<td>0.51 (0.10 to 0.92)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$I^2$</td>
<td>98.5</td>
<td>54.2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2 (df)$</td>
<td>332.71 (2)</td>
<td>4.37 (2)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>.11</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>–1.41 (–4.02 to 1.20)</td>
<td>2.56 (–1.22 to 6.35)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$I^2$</td>
<td>99</td>
<td>98.2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2 (df)$</td>
<td>1.06 (1)</td>
<td>55.43 (1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$P$ value</td>
<td>.29</td>
<td>&lt;.001</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Family member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS-depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>–0.25 (–0.72 to 0.21)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$I^2$</td>
<td>0</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2 (df)$</td>
<td>0.41 (1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$P$ value</td>
<td>.52</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS-anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled ES, Hedges g (95% CI)</td>
<td>–0.23 (–0.70 to 0.23)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$I^2$</td>
<td>0</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity, $\chi^2 (df)$</td>
<td>0.65 (1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>
Quality of Life

Quality of life was measured by the Functional Assessment of Cancer Therapy–Breast [44,54,80,93,95], Functional Assessment of Cancer Therapy–General [38,53,57,86,88], European Organization for Research and Treatment of Cancer Quality of Life Questionnaire, 30-items [35,40,43,74,91,99], and 36-item Short Form Survey [41,54,64,70,71]. Overall, participants receiving interventions with 338 participants in the intervention groups (without subscale scores reported) were reported in 5 studies. Overall, interventions were not more effective than control conditions for reducing anxiety (Hedges \(g\) = –0.12, 95% CI –0.19 to 0.43), with high heterogeneity of 90.2% (\(P<.001\)). With respect to publication bias, the funnel plot displayed a greater number of studies toward the top of the mean (Egger test, \(P<.001\)). The statistical heterogeneity among studies was \(I^2=0\%\) across all time-varying effects.

### Anxiety and Depression

Anxiety was assessed by the HADS-anxiety subscale [57,64,69,74,86,95,96,99], State-Trait Anxiety Inventory (STAI) [66,71,72,82,83], and a combination of the STAI and HADS-anxiety in 1075 participants in the intervention groups. Overall, interventions were not more effective than control conditions for reducing anxiety (Hedges \(g\) = 0.01, 95% CI –0.10 to 0.16), with a high heterogeneity of 60.9% (\(P<.001\)). With respect to publication bias, the funnel plot displayed a greater number of studies toward the top of the mean (Egger test, \(P<.001\)). The statistical heterogeneity among studies was \(I^2=69.4\%\) for the immediate effect and \(I^2=29.8\%\) for the interim effect.

### Depression

Depression was assessed by the HADS-depression subscale [50,69,74,86,95,96,99], Center for Epidemiologic Studies Depression Scale [41,68,71,77], Patient Health Questionnaire-9 (PHQ-9) [40,53], and a combination of the PHQ-9 and HADS-anxiety [43,57] in 1509 participants in the intervention groups. Overall, interventions were not more effective than control conditions for reducing depression (Hedges \(g\) = 0.03, 95% CI –0.10 to 0.16), with a high heterogeneity of 60.9% (\(P<.001\)). With respect to publication bias, the funnel plot displayed a greater number of studies toward the top of the mean (Egger test, \(P<.001\)). The statistical heterogeneity among studies was \(I^2=69.4\%\) for the immediate effect and \(I^2=29.8\%\) for the interim effect.

<table>
<thead>
<tr>
<th>Population, outcome, measure, and value</th>
<th>Effect at different time points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall(^a)</td>
<td>Immediate</td>
</tr>
<tr>
<td>(P) value</td>
<td>.42</td>
</tr>
</tbody>
</table>

\(^a\)The overall effect accounts for time-varying effect across different time points.

\(^b\)QOL: quality of life.

\(^c\)FACT-B: Functional Assessment of Cancer Therapy–Breast.

\(^d\)ES: effect size.

\(^e\)Not applicable.


\(^g\)QLQ-30: Quality of Life Questionnaire, 30 items.

\(^h\)SF36: Short Form Survey 36-item.

\(^i\)HADS: Hospital Anxiety and Depression Scale.

\(^j\)CESD: Center for Epidemiologic Studies Depression Scale.

\(^k\)PHQ: Patient Health Questionnaire-9.

\(^l\)SATI: State-Trait Anxiety Inventory.

\(^m\)CT: Distress Thermometer.

\(^n\)CBI: Coping Behaviors Inventory.

\(^o\)ES: effect size.

\(^p\)QOL: quality of life.

\(^q\)FACT: Functional Assessment of Cancer Therapy.

\(^r\)PHQ-9: Patient Health Questionnaire-9.

\(^s\)SF36: Short Form Survey 36-item.

\(^t\)QLQ-30: Quality of Life Questionnaire, 30 items.

\(^u\)DT: Distress Thermometer.

\(^v\)CBI: Coping Behaviors Inventory.

\(^w\)FACT-G: Functional Assessment of Cancer Therapy–General.

\(^x\)FACT-B: Functional Assessment of Cancer Therapy–Breast.

\(^y\)PHQ-9: Patient Health Questionnaire-9.

\(^z\)SATI: State-Trait Anxiety Inventory.

\(^{+}\)CESD: Center for Epidemiologic Studies Depression Scale.

\(^{++}\)HADS: Hospital Anxiety and Depression Scale.
Distress
Psychological distress was assessed in 182 participants in the intervention groups using the distress thermometer [35,69,91]. Overall, participants in the intervention groups showed no reduction in distress, with a mean difference between groups of Hedges $g=0.98$ (95% CI −0.18 to 2.14). The impact of heterogeneity within the studies was significant ($I^2=98.5\%$; $P<.001$). Regarding publication bias, the funnel plot displayed a symmetric distribution around the mean effect (Egger test, $P=.46$). The immediate effect was Hedges $g=0.51$ (95% CI 0.10-0.92), with statistical heterogeneity $I^2=54.2\%$.

Self-Efficacy
Self-efficacy was measured by the Coping Behaviors Inventory in 174 participants in the intervention groups [44,55]. Overall, participants in the intervention groups did not report improvement in self-efficacy, with a standardized mean difference of Hedges $g=-1.41$ (95% CI −4.02 to 1.20). However, the impact of heterogeneity within studies was significant ($I^2=99\%$; $P<.001$). Regarding publication bias, the funnel plot displayed a symmetric distribution around the mean effect (Egger test, $P=.22$). The immediate effect was Hedges $g=2.56$ (95% CI −1.22 to 6.35) with high heterogeneity ($I^2=98.2\%$; $P<.001$).

Subgroup Analyses
Given the heterogeneity of reporting on dosage information and limited data, the subgroup analysis of dosage on intervention effect was not conducted. Table 2 includes the results of the subgroup analysis on the effect on quality of life, depression and anxiety, and distress. Overall, the associations between delivery mode and control condition with patient outcomes were not statistically significant ($P>.05$).

Table 2. Subgroup analyses on the effect of delivery mode (internet vs noninternet) and control condition (usual care vs active control) on patient outcomes.

<table>
<thead>
<tr>
<th>Outcome and moderators</th>
<th>Effect size, Hedges $g$ (95% CI)</th>
<th>SE</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of life (27 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery mode</td>
<td>0.04 (−0.06 to 0.14)</td>
<td>0.05</td>
<td>.45</td>
</tr>
<tr>
<td>Control condition</td>
<td>−0.01 (−0.99 to 0.06)</td>
<td>0.04</td>
<td>.78</td>
</tr>
<tr>
<td>HADS$^a$ total (6 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery mode</td>
<td>0.16 (−0.30 to 0.62)</td>
<td>0.24</td>
<td>.50</td>
</tr>
<tr>
<td>Control condition</td>
<td>N/A$^b$</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Depression (21 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery mode</td>
<td>0.10 (−0.09 to 0.29)</td>
<td>0.10</td>
<td>.31</td>
</tr>
<tr>
<td>Control condition</td>
<td>−0.04 (−0.17 to 0.09)</td>
<td>0.07</td>
<td>.55</td>
</tr>
<tr>
<td>Anxiety (15 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery mode</td>
<td>0.01 (−0.07 to 0.09)</td>
<td>0.04</td>
<td>.79</td>
</tr>
<tr>
<td>Control condition</td>
<td>−0.06 (−0.34 to 0.22)</td>
<td>0.14</td>
<td>.67</td>
</tr>
<tr>
<td>Distress (8 studies)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery mode</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Control condition</td>
<td>−0.03 (−0.34 to 0.28)</td>
<td>0.15</td>
<td>.86</td>
</tr>
</tbody>
</table>

$^a$HADS: Hospital Anxiety and Depression Scale.
$^b$N/A: not applicable.

Family Members’ Outcomes
The forest plots for overall effect sizes are displayed in Multimedia Appendix 5. The funnel plots for overall effect sizes are displayed in Multimedia Appendix 6. For family members’ data, we pooled 2 studies [36,69] on anxiety and depression for the meta-analysis with 68 participants in the intervention groups. Due to the small sample size, no time-varying effect or subgroup analysis was calculated. The overall effect on anxiety was Hedges $g=−0.23$ (95% CI −0.70 to 0.23), with heterogeneity of $I^2=0\%$ ($P=42$). The overall effect on depression was Hedges $g=−0.25$ (95% CI −0.72 to 0.21), with heterogeneity of $I^2=0\%$ ($P=.52$). Regarding publication bias, the funnel plot displayed asymmetrical scattered points with statistical significance (Egger test, $P<.001$).

Discussion
Overview
This systematic review and meta-analysis of 65 unique digital psychosocial intervention studies for patients with cancer and their family members provides strong evidence that psychosocial interventions delivered through digital health significantly improve psychosocial outcomes. There were 3 major findings. First, this review included a large group of participants with various cancer diagnoses; however, underrepresented populations affected by cancer were not included, and the results...
predominantly focused on White patients. Second, we found that various intervention modes and components were used. There is a lack of specificity with respect to the description of interventions or theoretical basis for interventions, which may hinder future replication or refinement of the interventions and understanding of underlying mechanisms. Third, despite high heterogeneity across studies, the available data suggest that digital psychosocial interventions effectively improve some psychosocial outcomes, including patients’ quality of life, anxiety, and depression.

**Principal Findings**

First, the majority of participants in the included studies were White and female, which does not reflect the broader patient population with cancer, including non-White ethno-racial groups (ie, African American or Black, American Indian and Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and Hispanic or Latino populations). It is well documented in the literature that the impact of cancer on psychological distress and quality of life is worse for racial and ethnic minority groups [100-102]. Therefore, future trials should include more participants from underrepresented groups to reduce health care disparities and improve generalizability in diverse populations [103]. Family members and caregivers were rarely included in the studies reviewed. However, there is ample evidence that family members and caregivers experience significant caregiver burden, worsening quality of life, and difficulty with psychological adjustment, therefore needing support [104,105]. Previous systematic reviews suggest that interventions targeting problem-solving and communication skills may ease the burden related to patient care and improve caregivers’ quality of life [106]. Many reviews focus on the evaluation of nondigital interventions targeting the psychosocial experience in family members and caregivers, including several reviews of caregiver interventions [9,107-109]. Therefore, with growing technology usage, more digital interventions are needed to address family members’ or caregivers’ needs.

Few RCTs met all quality criteria, including blinding, analysis by treatment assignment, and standardized outcome assessment [110]. While concealing assignments from participants and those delivering interventions is not always possible, single blinding of assessors should occur in well-designed research. Few studies used power calculations for sample size, making it difficult to determine whether sample sizes were adequate [111]. Generally, results from group sizes <20 are questionable. There are several effective strategies known to increase the retention rate, such as adding monetary incentives and using an open trial design [112,113]. The critical appraisal also depends on comprehensive reporting of study details, which were limited in the identified studies. Although attempts have been made to improve reporting using the CONSORT (Consolidated Standards of Reporting Trials) statement for RCTs and the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) statement for nonrandomized intervention studies in the early 2000s [114,115], we did not see improvement in appraisal scores over time. The main limitations of the results include underpowered and methodologically weaker studies. These highlight the need for improved methodologies in future research, as the overall methodological quality was moderate.

Second, this study identified various intervention modes and components, which is consistent with a previous systematic review of psychosocial interventions for patients with advanced cancer, which identified similar intervention components, including psychoeducation and CBT-based intervention, as 2 of the most common [11]. However, more than half of the included studies did not use theoretical or conceptual frameworks to guide the development of intervention components or selection of outcomes. The lack of a theoretical framework leads to a lack of clarity about the mechanisms through which intervention components impact psychosocial outcomes [116]. In future research, theories or conceptual frameworks need to be incorporated to help us better understand the mechanisms that explain the changes in psychosocial outcomes when using digital health interventions.

In addition, the prescribed dosage information (ie, the number of sessions, duration, and frequency) was inconsistently reported, making it difficult to estimate an efficacious intervention dose. Most of the interventions were self-paced, without the involvement of an interventionist, which gives the patient autonomy to choose which intervention component or module they would like to focus on and how much time to allocate. However, there is a lack of information on intervention uptake, which may have influenced the effectiveness of interventions. Approaches that tackle barriers to adherence at various levels (eg, individual, family, clinician, agency, and environment) and improve engagement should be implemented [117]. For example, a scoping review about engagement strategies in digital interventions for mental health promotion recommended personalized feedback, e-coaching to guide content and individual progress, social platforms and interaction with peers, content gamification, reminders, and ease of use [118].

Third, we found some significant improvement in the patient’s quality of life. Some studies with a smaller number of participants or with a focus on internet-based interventions reported an improvement, but the results from these studies were not consistent [21,119]. This meta-analysis, including 21 studies, revealed a small effect size for overall effectiveness of digital health interventions in improving patients’ quality of life (Hedges g=0.05, 95% CI –0.01 to 0.10), with time-varying effects shown as promising. Another meta-analysis that pooled 16 studies demonstrated a larger positive effect of mHealth interventions on the quality of life of patients with cancer (standardized mean difference 0.28, 95% CI 0.03-0.53) [21]. Another meta-analysis that included 6 internet-based psychoeducational interventions for patients with cancer showed no significant improvement in quality of life (mean difference 1.10, 95% CI –4.42 to 6.63) [119]. Importantly, our analyses found the largest improvements in quality of life occurred from post intervention to 8 weeks (Hedges g=2.25, 95% CI 0.36-4.14). The effect of psychosocial interventions decreased after 8 weeks of follow-up, suggesting that interventions may need booster sessions or tailoring to time-sensitive needs in order to maintain effectiveness in the long term. This result was limited by substantial inconsistency across studies in all evaluation periods except the medium-term effect [32]. In addition, given the heterogeneity of follow-up periods in selected
studies, the time-varying effect was only tested with a small number of studies, not in the 21 studies we used to calculate the overall effect.

This meta-analysis was able to demonstrate the effectiveness of digital health interventions on both anxiety and depression (measured by HADS: Hedges $g=-0.72$, 95% CI $-1.89$ to $0.46$). Our finding was partially consistent with the other meta-analyses. One study showed that internet-based psychoeducational interventions had a significant effect on decreasing depression (standardized mean difference $-0.58$, 95% CI $-1.12$ to $-0.03$), but found no evidence for effects on distress (standardized mean difference $-1.03$, 95% CI $-2.63$ to $0.57$) [119]. However, there was considerable heterogeneity in measurements among the studies included in the review by Wang et al [119]; it is difficult to determine how meaningful it is to make direct comparisons between the studies included in this meta-analysis and past reports. Possible ways to address this problem could be using similar outcome measures and a standardized study report.

Our meta-analysis demonstrated that the interventions were effective in reducing anxiety and depression in family caregivers. However, the effect size was small, perhaps due to the limited number of studies. This is partially consistent with findings from another meta-analysis, which found depressive symptoms decreased from baseline to post intervention (Hedges $g=-0.44$, 95% CI $-1.03$ to $0.15$) [120], while anxiety remained relatively stable when comparing intervention to control either at postintervention (Hedges $g=0.12$, 95% CI $-0.16$ to $0.44$) or during follow-up (Hedges $g=-0.08$, 95% CI $-0.34$ to $0.19$).

Strengths and Limitations
This review's strength lies in its rigorous design, sophisticated data synthesis, and enduring empirical contributions. We acknowledge that our literature search was conducted 4 years before manuscript submission. The findings and contributions from our research remain pertinent and enduring. This is because, as digital psychosocial interventions continue to evolve, the core intervention content and outcomes have remained relatively consistent over the past 4 years. It would be valuable to conduct a reassessment of the evolving body of evidence concerning digital psychosocial interventions that have emerged since the onset of the COVID-19 pandemic. Moreover, this study was conducted in line with best practices by double-coding and following the PRISMA guidelines. The meta-analysis, including subgroup analysis, was conducted using appropriate methods for combining studies across various follow-up periods. Although we did an extensive search at the start of this review, we may have missed some critical studies, unreported, or unfinished studies. If all data were available, the meta-analysis could have reduced the chances of inflated type-1 error for both observed and unobserved effects that were available for assessment [121]. There was not enough data to perform post hoc analyses to examine the effect of factors such as intervention components and length of intervention on outcomes due to insufficient data.

Conclusions
Patients with cancer and their family members need high-quality psychosocial interventions throughout the cancer trajectory. Digital technologies provide a platform to deliver evidence-based psychosocial interventions from a distance, without the heightened risk of contracting viruses, especially for patients with cancer whose immune systems are compromised. This study comprehensively synthesized the effects of digital psychosocial interventions for people affected by cancer. Our findings suggest that digital health interventions are effective for adult patients with cancer and their family members. Further research development in this area needs to include large, high-quality studies with a clear description of the methodology, theoretical foundations, and standardized tools to permit inclusion in meta-analyses to inform the effectiveness of interventions for a better understanding of the mechanisms.

Acknowledgments
Dr Flannery receives salary support from the National Institute of Health National Cancer Institute (grants U01 CA233167 and UG1 CA189961).

Conflicts of Interest
None declared.

Multimedia Appendix 1
Literature search strategy.
[DOCX File , 16 KB - cancer_v10i1e46116_app1.docx ]

Multimedia Appendix 2
Summary tables.
[DOCX File , 53 KB - cancer_v10i1e46116_app2.docx ]

Multimedia Appendix 3
Forest plots of studies reported on intervention effect on patient outcomes.
[DOCX File , 1723 KB - cancer_v10i1e46116_app3.docx ]
References


76. Mohammad EB, Ahmad M. Virtual reality as a distraction technique for pain and anxiety among patients with breast cancer: a randomized control trial. Palliat Support Care 2019;17(1):29-34. [FREE Full text] [doi: 10.1017/S1478951518000639] [Medline: 30198451]


Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
HADS: Hospital Anxiety and Depression Scale
JBI: Joanna Briggs Institute
PHQ-9: Patient Health Questionnaire-9
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
STAI: State-Trait Anxiety Inventory
TREND: Transparent Reporting of Evaluations with Nonrandomized Designs

© Yingzi Zhang, Marie Flannery, Zhihong Zhang, Meghan Underhill-Blazey, Melanie Bobry, Natalie Leblanc, Darcey Rodriguez, Chen Zhang. Originally published in JMIR Cancer (https://cancer.jmir.org), 05.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

Please cite as:
Digital Health Psychosocial Intervention in Adult Patients With Cancer and Their Families: Systematic Review and Meta-Analysis
JMIR Cancer 2024;10:e46116
URL: https://cancer.jmir.org/2024/1/e46116
doi: 10.2196/46116
PMID: 38315546
Abstract

Background: People with cancer frequently experience severe and distressing symptoms associated with cancer and its treatments. Predicting symptoms in patients with cancer continues to be a significant challenge for both clinicians and researchers. The rapid evolution of machine learning (ML) highlights the need for a current systematic review to improve cancer symptom prediction.

Objective: This systematic review aims to synthesize the literature that has used ML algorithms to predict the development of cancer symptoms and to identify the predictors of these symptoms. This is essential for integrating new developments and identifying gaps in existing literature.

Methods: We conducted this systematic review in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist. We conducted a systematic search of CINAHL, Embase, and PubMed for English records published from 1984 to August 11, 2023, using the following search terms: cancer, neoplasm, specific symptoms, neural networks, machine learning, specific algorithm names, and deep learning. All records that met the eligibility criteria were individually reviewed by 2 coauthors, and key findings were extracted and synthesized. We focused on studies using ML algorithms to predict cancer symptoms, excluding nonhuman research, technical reports, reviews, book chapters, conference proceedings, and inaccessible full texts.

Results: A total of 42 studies were included, the majority of which were published after 2017. Most studies were conducted in North America (18/42, 43%) and Asia (16/42, 38%). The sample sizes in most studies (27/42, 64%) typically ranged from 100 to 1000 participants. The most prevalent category of algorithms was supervised ML, accounting for 39 (93%) of the 42 studies. Each of the methods—deep learning, ensemble classifiers, and unsupervised ML—constituted 3 (3%) of the 42 studies. The ML algorithms with the best performance were logistic regression (9/42, 17%), random forest (7/42, 13%), artificial neural networks (5/42, 9%), and decision trees (5/42, 9%). The most commonly included primary cancer sites were the head and neck (9/42, 22%) and breast (8/42, 19%), with 17 (41%) of the 42 studies not specifying the site. The most frequently studied symptoms were xerostomia (9/42, 14%), depression (8/42, 13%), pain (8/42, 13%), and fatigue (6/42, 10%). The significant predictors were age, gender, treatment type, treatment number, cancer site, cancer stage, chemotherapy, radiotherapy, chronic diseases, comorbidities, physical factors, and psychological factors.

Conclusions: This review outlines the algorithms used for predicting symptoms in individuals with cancer. Given the diversity of symptoms people with cancer experience, analytic approaches that can handle complex and nonlinear relationships are critical. This knowledge can pave the way for crafting algorithms tailored to a specific symptom. In addition, to improve prediction precision, future research should compare cutting-edge ML strategies such as deep learning and ensemble methods with traditional statistical models.
machine learning; ML; deep learning; DL; cancer symptoms; prediction model

Introduction

Background

Cancer poses considerable physical and psychological challenges for those diagnosed with the disease. The Global Cancer Observatory estimated that there were 19.3 million new cancer cases and 43.8 million individuals living with cancer within 5 years of diagnosis globally in 2020 [1]. Symptoms such as fatigue, pain, nausea, vomiting, depression, and anxiety often persist beyond treatment [2-5], detrimentally affecting individuals’ quality of life [6]. Moreover, people with cancer frequently grapple with multiple intertwined symptoms [7], intensifying their distress [8]. Unmanaged cancer symptoms can lead to increased health care use, including emergency department visits and unscheduled hospitalizations to address these symptoms; a decline in the quality of life [9]; and even a reduced life expectancy. Providing precision symptom management tailored to the individual at the right moment has the potential to significantly improve outcomes, which is crucial for both people with cancer and their health care providers. Accurately predicting and addressing these symptoms is fundamental to providing such precision in symptom management.

Artificial intelligence, incorporating machine learning (ML) and deep learning (DL) models, excels in handling complex, high-dimensional, and noisy data. It has demonstrated effectiveness in disease diagnosis, predicting disease recurrence, enhancing quality of life, and symptom management [10-16]. There is a growing interest in ML in the emerging field of predictive analytics for cancer symptoms. ML contributes to the development of robust clinical decision systems, enhancing overall health care delivery [17]. ML algorithms can be broadly categorized into supervised learning, unsupervised learning, semisupervised learning, and reinforcement learning. DL, a subset of ML, addresses complex tasks such as speech recognition, image identification, and natural language processing [18].

Objectives

This study seeks to offer a comprehensive and systematic review of the literature on the application of ML algorithms in predicting symptoms for people with cancer. Conducting this review of a rapidly expanding body of literature is imperative to understand the current state of the science for ML models in symptom prediction for cancer and to guide future research. This research aims to provide a comprehensive understanding of the current state of research; identify areas for improvement; and understand the limitations and gaps in the current literature, such as a lack of specific focus on ML models for patients with cancer. By comparing model performances across diverse symptom prediction tasks, we can identify the best practices, highlight areas for improvement, and offer informed recommendations that will propel the field of predictive analytics in cancer symptom research forward.

Methods

Search Strategy and Data Sources

This study was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) protocol [19] and involved a comprehensive database search spanning from 1984 to August 11, 2023, including the PubMed, Embase, CINAHL, and Google Scholar databases. The search terms encompassed cancer, neoplasm, signs and symptoms, neural networks, machine learning, and specific algorithm names. In our study, we used Boolean expressions, using specific combinations of keywords and phrases, acknowledging the variability in terminology across studies. Search results were compiled using EndNote 20 (Clarivate Analytics). The detailed search strategy, which uses Boolean expressions, and the PRISMA checklist can be found in Multimedia Appendices 1 and 2.

Inclusion and Exclusion Criteria

To identify relevant research focusing on the application of ML methods in predicting cancer symptoms, we applied the following inclusion criteria: (1) papers published in English, (2) studies that used ML algorithms, and (3) research specifically aimed at predicting cancer symptoms. The exclusion criteria were as follows: (1) nonhuman studies, (2) technical reports, (3) review papers, (4) book chapters or series, (5) conference proceedings, and (6) studies for which full texts were unavailable. Two authors, NZ and NY, independently screened and cross-checked the candidate records. During the screening process, conducted using EndNote 20, any disagreements were resolved by consulting a third reviewer (SGW). The screening process involved an initial review of titles and abstracts, followed by a full-text examination to determine the study’s eligibility for inclusion in the review.

Data Extraction and Analysis

In our study, we implemented a systematic, multistep process for data synthesis. Initially, relevant studies were identified and selected based on the predefined inclusion and exclusion criteria. Two independent researchers, NZ and NY, extracted data from 42 selected studies. They worked independently to mitigate bias and enhance the accuracy of the data extraction process. In cases of discrepancies, these were resolved through discussion or consultation with a third reviewer, SGW. The extracted data were aggregated, involving the collation of study characteristics such as research location, sample size, study design, types of ML algorithms, validation metrics, identified significant predictors, cancer types, and the specific symptoms focused on. This comprehensive approach enabled us to reduce the bias and increase the reliability of our findings. For the analysis, we used both quantitative and qualitative methods. Quantitative data,
such as frequencies and percentages, were compiled and analyzed using Python. This included the creation of insightful plots and heat maps to identify patterns and trends, illustrating relationships among variables and highlighting key findings in an easily digestible format. Qualitative aspects, such as algorithm implementation or study design, were explored through narrative synthesis. This allowed for a deeper understanding of the context and nuances in the application of ML algorithms for cancer symptom prediction.

We conducted a cross-analysis to compare findings from different studies, assessing the effectiveness of various ML algorithms across different cancer types and symptoms and identifying common predictors of success and the challenges faced. Finally, we interpreted the findings in the context of the existing literature. We discussed how our results align with or differ from previous studies and what new insights our synthesis brings to the field of ML in cancer symptom prediction.

### Results

#### Overall Results

A search across the 3 databases produced 1788 papers. After removing 289 duplicates, we screened the records for titles and abstracts, excluding another 1352 irrelevant records. However, 1 study was not retrieved. We reviewed the full text of the remaining 146 records, omitting 105 due to the absence of ML application in predicting cancer symptoms (69/146, 47.3%), not being a research article (34/146, 23.3%), and not being an English article (1/146, 1%). In the second phase, we intend to include Google Scholar in our research methodology to capture an additional 113 articles not found in our main databases, although 1 study was not retrieved. We reviewed the full text of the remaining 99 records, ultimately excluding all of them for reasons such as the lack of ML applications in cancer symptom prediction (89/99, 90%) and not being a research article (10/99, 10%). Eventually, 42 studies met the inclusion criteria, as depicted in Figure 1.

![PRISMA Flowchart](https://cancer.jmir.org/2024/1/e52322/有多么图片.png)

Of the 42 studies, 42 (100%) is listed in PubMed, Embase covers 37 (88%) studies, and CINAHL includes 18 (43%) studies. The distribution and overlap of these research articles across the databases are illustrated in Multimedia Appendix 3.

The data extracted from these studies, which include the reference number, research location, year, data type, cancer site, symptoms, significant predictors, ML algorithms, and validation methods, are detailed in Table 1 and in Multimedia Appendix 4.

---

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. ML: machine learning.
Table 1. Details of the included studies (n=42).

<table>
<thead>
<tr>
<th>Study</th>
<th>Country, year</th>
<th>Data type; number of data</th>
<th>Population</th>
<th>Cancer symptoms</th>
<th>Significant predictors</th>
<th>Algorithms</th>
<th>Validation methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun et al [20]</td>
<td>China, 2023</td>
<td>Clinical data; 1152</td>
<td>People with breast cancer</td>
<td>Pain</td>
<td>Postmenopausal status, urban medical insurance, history of at least one operation, underwent general anesthesia with fentanyl and sevoflurane, and received axillary lymph node dissection.</td>
<td>LR(^{a,b}), RF(^c), GB-DT(^d), and XGB(^e)</td>
<td>Random</td>
</tr>
<tr>
<td>Xinran et al [21]</td>
<td>China, 2023</td>
<td>Clinical data; 494</td>
<td>People with advanced cancer</td>
<td>Cognitive impairment</td>
<td>Cancer course, anxiety, and age</td>
<td>LR and ANN(^f)</td>
<td>Random</td>
</tr>
<tr>
<td>Shaikh et al [22]</td>
<td>United States, 2023</td>
<td>Clinical data; 1152</td>
<td>Survivors of cancer with osteoarthritis</td>
<td>Depression</td>
<td>Age, education, care fragmentation, polypharmacy, and zip code–level poverty</td>
<td>XGB</td>
<td>10-fold CV(^g)</td>
</tr>
<tr>
<td>Kober et al [23]</td>
<td>United States, 2023</td>
<td>Clinical data; 1217</td>
<td>People with cancer receiving chemotherapy</td>
<td>Morning fatigue</td>
<td>13 individual Li-Fraumeni syndrome items</td>
<td>EN(^h), RF, LASSO(^i), LR (filtered/unfiltered), RPAIR(^j), and SVM(^k)</td>
<td>Random</td>
</tr>
<tr>
<td>Du et al [24]</td>
<td>China, 2023</td>
<td>Clinical data; 565</td>
<td>People with cancer</td>
<td>Fatigue</td>
<td>Pain score, Eastern Cooperative Oncology Group score, platelet distribution width, and continuous erythropoiesis receptor activator</td>
<td>L.R, RF, NB(^l), and XGB</td>
<td>5-fold CV</td>
</tr>
<tr>
<td>Moscato et al [25]</td>
<td>Italy, 2022</td>
<td>Clinical data; 21</td>
<td>People with cancer</td>
<td>Pain</td>
<td>N/A(^m)</td>
<td>SVM, RF, MP(^n), LR, and AdaBoost(^o)</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Masukawa et al [26]</td>
<td>Japan, 2022</td>
<td>Clinical data; 808</td>
<td>People with cancer</td>
<td>Social distress, spiritual pain, pain, dyspnea, nausea, and insomnia</td>
<td>N/A</td>
<td>LR, RF, light GBM(^p), SVM, and ensemble</td>
<td>5-fold CV</td>
</tr>
<tr>
<td>Fanizzi et al [27]</td>
<td>Italy, 2022</td>
<td>CT(^q) image data; 61</td>
<td>People with oropharyngeal cancer receiving radiotherapy</td>
<td>Xerostomia</td>
<td>Weight preradiotherapy, induction chemotherapy, sex, platinum-based chemotherapy, current chemotherapy, alcohol history, age at diagnosis, smoking history, surgery, clinical tumor, and clinical node</td>
<td>SVM and CNN(^r)</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Ueno et al [28]</td>
<td>Japan, 2022</td>
<td>Clinical data; 284</td>
<td>People with breast cancer</td>
<td>Insomnia</td>
<td>General fatigue, physical fatigue, and cognitive fatigue</td>
<td>L2 penalized LR and XGB</td>
<td>8-fold CV</td>
</tr>
<tr>
<td>On et al [29]</td>
<td>Korea, 2022</td>
<td>Clinical data; 935</td>
<td>People with cancer receiving chemotherapy</td>
<td>Nausea-vomiting, fatigue-anorexia, diarrhea, hypersensitivity, stomatitis, hand-foot syndrome, peripheral neuropathy, and constipation</td>
<td>L.R, DT(^s), and ANN</td>
<td>3-fold CV</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country, year</td>
<td>Data type; number of data</td>
<td>Population</td>
<td>Cancer symptoms</td>
<td>Significant predictors</td>
<td>Algorithms</td>
<td>Validation methods</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>----------------------------</td>
<td>---------------------------------------------------</td>
<td>-----------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Li et al [30]</td>
<td>China, 2022</td>
<td>Clinical data and CT image data; 365</td>
<td>People with cancer receiving radiotherapy</td>
<td>Xerostomia</td>
<td>Hypertension, age, total radiotherapy dose, dose at 50% of the left parotid volume, mean dose to right parotid gland, mean dose to oral cavity, and course of induction chemotherapy</td>
<td>RF, DT and XGB</td>
<td>External validation</td>
</tr>
<tr>
<td>Kurisu et al [31]</td>
<td>Japan, 2022</td>
<td>Clinical data; 668</td>
<td>People with advanced cancer receiving pharmacological interventions</td>
<td>Delirium</td>
<td>The baseline Delirium Rating Scale-R98 severity score (cutoff of 15), hypoxia, and dehydration</td>
<td>DT</td>
<td>5-fold CV</td>
</tr>
<tr>
<td>Guo et al [32]</td>
<td>China, 2022</td>
<td>Clinical data; 80</td>
<td>People with lung cancer receiving chemotherapy</td>
<td>Lung infection</td>
<td>Age ≥60 years, length of stay ≥14 days, surgery history, combined chemotherapy, myelosuppression, diabetes, and hormone application</td>
<td>LR and ANN</td>
<td>Random</td>
</tr>
<tr>
<td>Baglione et al [33]</td>
<td>United States, 2022</td>
<td>Clinical data; 40</td>
<td>People with breast cancer</td>
<td>Depressed mood and anxiety</td>
<td>Connectedness, receive support, frequency and duration use of mobile app, and physical pain</td>
<td>RF and XGB</td>
<td>LOOCV</td>
</tr>
<tr>
<td>Chao et al [34]</td>
<td>United States, 2022</td>
<td>Clinical data and CT image data; 155</td>
<td>People with HNC receiving radiotherapy</td>
<td>Xerostomia</td>
<td>N/A</td>
<td>SVM, KNN, NB, and RF</td>
<td>Nested</td>
</tr>
<tr>
<td>Wakabayashi et al [35]</td>
<td>Japan, 2021</td>
<td>Clinical data and CT image data; 69</td>
<td>People with cancer receiving radiotherapy</td>
<td>Pain</td>
<td>Age, numeric rating scale, and biological effective dose 10</td>
<td>RF</td>
<td>LOOCV</td>
</tr>
<tr>
<td>Zhou et al [36]</td>
<td>China, 2021</td>
<td>Clinical data; 386</td>
<td>People with colorectal cancer after chemotherapy</td>
<td>Cognitive impairment</td>
<td>Age, BMI, colostomy, treatment complications, cancer-related anemia, depression, diabetes, Quality of Life Questionnaire Core 30 score, exercise, hypercholesterolemia, diet, marital status, education level, and pathological stage</td>
<td>RF, LR, and SVM</td>
<td>Random</td>
</tr>
<tr>
<td>Xuyi et al [37]</td>
<td>Canada, 2021</td>
<td>Clinical data; 46,104</td>
<td>Specific cancer site or treatment not mentioned</td>
<td>Pain, depression, and well-being</td>
<td>Lung cancer, late-stage cancer, existing chronic conditions such as osteoarthritis, mood disorder, hypertension, diabetes, and coronary disease</td>
<td>ANN</td>
<td>Random</td>
</tr>
<tr>
<td>Xu et al [38]</td>
<td>China, 2021</td>
<td>Clinical data; 598</td>
<td>People with gastrointestinal tumors after surgery</td>
<td>Postoperative fatigue</td>
<td>Age, higher degree of education, lower personal monthly income, advanced cancer, hypoproteinemia, preoperative anxiety or depression, and limited social support</td>
<td>LR, ANN, CART</td>
<td>Random</td>
</tr>
<tr>
<td>Wei et al [39]</td>
<td>China, 2021</td>
<td>Clinical data; 533</td>
<td>People with breast cancer</td>
<td>Lymphedema</td>
<td>N/A</td>
<td>ANN, LR, C5.0, RF, SVM, CART</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Wang et al [40]</td>
<td>United States, 2021</td>
<td>Clinical data; 823</td>
<td>People with HNC</td>
<td>Pain, taste, and general activity</td>
<td>N/A</td>
<td>SVM, KNN, and RF; Gaussian NB and MLP; and ARIMA and LSTM</td>
<td>Random</td>
</tr>
<tr>
<td>Study</td>
<td>Country, year</td>
<td>Data type; number of data</td>
<td>Population</td>
<td>Cancer symptoms</td>
<td>Significant predictors</td>
<td>Algorithms</td>
<td>Validation methods</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------</td>
<td>--------------------------------------------</td>
<td>-----------------------------------------------------</td>
<td>--------------------------------------</td>
<td>------------------------</td>
<td>-------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Wang et al [41]</td>
<td>United States, 2021</td>
<td>Clinical data and CT image data; 138</td>
<td>Specific cancer site or treatment not mentioned</td>
<td>Depression</td>
<td>N/A</td>
<td>Fine tree, medium tree, coarse tree, linear-discriminant, quadratic discriminant, LR, Gaussian NB, kernel NB, linear SVM, quadratic SVM, cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM, Fine KNN, Medium KNN, Coarse KNN, Cubic KNN, Weighted KNN, boosted trees, bagged trees, subspace discriminant, subspace KNN, and random undersampling boosted trees</td>
<td>5-fold CV</td>
</tr>
<tr>
<td>Mosa et al [17]</td>
<td>United States, 2021</td>
<td>Clinical data; 6124</td>
<td>People with cancer receiving chemotherapy</td>
<td>Nausea-vomiting</td>
<td>Smoking, alcohol status, sex, age, and BMI</td>
<td>NB, LR, ANN, SVR&lt;sup&gt;ad&lt;/sup&gt;, and DT</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Low et al [42]</td>
<td>United States, 2021</td>
<td>Clinical data; 44</td>
<td>People with pancreatic cancer after surgery</td>
<td>Diarrhea, fatigue, and pain</td>
<td>Physical activity bouts, sleep, heart rate, and location</td>
<td>LR, KNN, SVM, RF, GB&lt;sup&gt;ab&lt;/sup&gt;, XGB, and LightGBM</td>
<td>3-fold CV and LOOCV</td>
</tr>
<tr>
<td>Kourou et al [43]</td>
<td>Greece, 2021</td>
<td>Clinical data; 609</td>
<td>People with breast cancer</td>
<td>Depression</td>
<td>A set of psychological traits (optimism, perceived ability to cope with trauma, resilience as a trait, and ability to understand the illness) and subjective perceptions of personal functionality (physical, social, and cognitive)</td>
<td>RF, SVM, and GB</td>
<td>5-fold CV</td>
</tr>
<tr>
<td>Kober et al [44]</td>
<td>United States, 2021</td>
<td>Clinical data; 1217</td>
<td>People with cancer receiving chemotherapy</td>
<td>Evening fatigue</td>
<td>Morning fatigue, lower evening energy, and sleep disturbance</td>
<td>RF, LR (filtered or unfiltered), RPAR, and SVM</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Hu et al [45]</td>
<td>China, 2021</td>
<td>Clinical data; 238</td>
<td>People with non-Hodgkin lymphoma receiving chemotherapy</td>
<td>Depression</td>
<td>Education level, sex, age, marital status, medical insurance, per capita monthly household income, pathological stage, Suicide Severity Rating Scale, Pittsburgh Sleep Quality Index, and Quality of Life Questionnaire Core 30</td>
<td>SVM, RF, and LASSO+LR</td>
<td>Random</td>
</tr>
<tr>
<td>Haun et al [46]</td>
<td>Germany, 2021</td>
<td>Clinical data; 496</td>
<td>People with cancer seen in primary care</td>
<td>Anxiety</td>
<td>Fatigue or weakness, insomnia, and pain appeared</td>
<td>OLS&lt;sup&gt;ad&lt;/sup&gt;, RR&lt;sup&gt;ad&lt;/sup&gt;, LASSO, ENR&lt;sup&gt;ad&lt;/sup&gt;, RF, and XGB</td>
<td>10-fold CV</td>
</tr>
</tbody>
</table>

*Note: CN indicates cancer, TM indicates tumor markers, CT indicates computed tomography.*
<table>
<thead>
<tr>
<th>Study</th>
<th>Country, year</th>
<th>Data type; number of data</th>
<th>Population</th>
<th>Cancer symptoms</th>
<th>Significant predictors</th>
<th>Algorithms</th>
<th>Validation methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al</td>
<td>United States, 2020</td>
<td>Clinical data and CT images data; 388</td>
<td>People with lung cancer after intensity-modulated radiation therapy</td>
<td>Weight loss</td>
<td>Joint Gross tumor volume L1+L2+L3 radiomics, Gross tumor volume, and esophageus L3 dosimetric</td>
<td>SVM, DNN[^4] and ensemble classifier</td>
<td>Nested CV</td>
</tr>
<tr>
<td>Juwara et al</td>
<td>Canada, 2020</td>
<td>Clinical data: 204</td>
<td>People with breast cancer after surgery</td>
<td>Anxiety, type of surgery, and acute pain</td>
<td>L$S_{2}^{ab}$, RR, ENR, RF, GB, and ANN</td>
<td>10-fold CV</td>
<td></td>
</tr>
<tr>
<td>Jiang et al</td>
<td>United States, 2019</td>
<td>Clinical data and CT images data; 427</td>
<td>People with HNC</td>
<td>Xerostomia</td>
<td>The patient has human papillomavirus, completed chemotherapy, their baseline xerostomia grade, tumor site, N stage, and use of feeding tube</td>
<td>RR, LASSO, and RF</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Sheikh et al</td>
<td>United States, 2019</td>
<td>CT images data; 266</td>
<td>People with HNC</td>
<td>Xerostomia</td>
<td>N/A</td>
<td>Generalized linear model</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Papachristou et al</td>
<td>United States, 2019</td>
<td>Clinical data: 799</td>
<td>People with cancer receiving chemotherapy</td>
<td>Sleep disturbance, anxiety, and depression</td>
<td>Age, gender, cancer site, the number of prior cancer treatment, and initial diagnosis</td>
<td>SVR (linear, polynomial, and radial Sigma) and n-CCA[^11]</td>
<td>10-fold CV and bootstrap</td>
</tr>
<tr>
<td>Zhang et al</td>
<td>China, 2018</td>
<td>Clinical data: 375</td>
<td>People with cancer receiving radiotherapy</td>
<td>Weight loss</td>
<td>Head and neck tumor location and total radiation dose of ≥70 Gray, and without postsurgery</td>
<td>DT and LR</td>
<td>Random</td>
</tr>
<tr>
<td>Olling et al</td>
<td>Denmark; 2018</td>
<td>Clinical and CT image; 131</td>
<td>People with lung cancer receiving radiotherapy</td>
<td>Odynophagia (painful swallowing)</td>
<td>N/A</td>
<td>Multivariable LR, Lasso and elastic net regularized general linear models, and SVM</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>Gabryś et al</td>
<td>Germany; 2018</td>
<td>Clinical and CT image: 153</td>
<td>People with HNC after radiotherapy</td>
<td>Xerostomia</td>
<td>The parotid gland volume, the spread of the contralateral dose-volume histogram, and the parotid gland eccentricity, and sex</td>
<td>LRL$<em>{1}^{bk}$, LRL$</em>{2}^{ad}$, LR-EN$<em>{2}^{am}$, KNN, SVM, ET$</em>{1}^{am}$, and GTB$_{10}^{mp}$</td>
<td>Single and nested CV</td>
</tr>
<tr>
<td>Lötisch et al</td>
<td>Germany; 2018</td>
<td>Clinical data: 1000</td>
<td>People with breast cancer after surgery</td>
<td>Pain</td>
<td>Age, chronic pain of any type, number of previous operations, BMI, preoperative pain in the area to be operated on, smoking and psychological factors</td>
<td>Unsupervised ML$_{10}^{mp}$</td>
<td>Random</td>
</tr>
<tr>
<td>Abdollahi et al</td>
<td>Iran; 2018</td>
<td>Clinical and CT image: 47</td>
<td>People with HNC receiving chemotherapy</td>
<td>Hearing loss</td>
<td>10 of the 490 radiomic features selected as the associated features with significant sensorineural hearing loss status</td>
<td>Decision stump, Hoefding, C4.5, NB, AdaBoost, bootstrap aggregating, and LR</td>
<td>10-fold CV</td>
</tr>
<tr>
<td>van Dijk et al</td>
<td>United States; 2018</td>
<td>Clinical data and CT image; 68</td>
<td>People with HNC</td>
<td>Xerostomia</td>
<td>N/A</td>
<td>LR</td>
<td>External validation</td>
</tr>
<tr>
<td>Cvetković</td>
<td>Serbia; 2017</td>
<td>Clinical data: 84</td>
<td>People with breast cancer</td>
<td>Depression</td>
<td>N/A</td>
<td>ELM$_{20}^{ol}$, ANN, and Fuzzy Genetic Algorithm</td>
<td>Random</td>
</tr>
</tbody>
</table>
A total of 2 individual researchers (NZ and NY) separately extracted data from each study, working independently of each other. This approach is used to reduce bias and increase the accuracy of the data extraction process. If discrepancies arise between the 2 independent authors, they are usually resolved through discussion or by consulting a third reviewer (SGW).

### Primary Database Information

The studies selected were published between 2017 and 2023 and were conducted in North America (18/42, 43%), Asia (16/42, 38%), and Europe (8/42, 19%). Methods of data collection varied, with studies originating from individual centers (23/42, 55%) and multiple centers (19/42, 45%).

<table>
<thead>
<tr>
<th>Study</th>
<th>Country, year</th>
<th>Data type; number of data</th>
<th>Population</th>
<th>Cancer symptoms</th>
<th>Significant predictors</th>
<th>Algorithms</th>
<th>Validation methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>van Dijk et al [60]</td>
<td>United States;2017</td>
<td>CT image features; 249</td>
<td>People with HNC</td>
<td>Xerostomia</td>
<td>N/A</td>
<td>LR</td>
<td>10-fold CV</td>
</tr>
</tbody>
</table>

*a*LR: logistic regression.

*Italic text in this column indicates the best results used in the study.*

*c*RF: random forest.

dGBDT: gradient boosting decision tree.

eXGB: extreme gradient boosting.

fANN: artificial neural network.

gCV: cross-validation.

hEN: elastic net.

iLASSO: Least absolute shrinkage and selection operator.

jRPAR: recursive partitioning and regression trees.

kSVM: support vector machine.

lNB: Naive bayes.

mN/A: not applicable.

nMP: multiple perceptron.

oAdaBoost: Adaptive boosting.

pGBM: light gradient boosting machine.

qCT: computed tomography.

rCNN: convolutional neural network.

sDT: decision tree.

tLOOCV: leave-one-out-cross-validation.

uHNC: head and neck cancer.

vKNN: k-nearest neighbor.

wCART: classification and regression tree.

xMLP: multilayer perceptron.

yARIMA: autoregressive integrated moving average.

zLSTM: long short-term memory neural network.

aaSVR: support vector regression.

abGB: gradient boosting.

acOLS: ordinary least square.

adRR: ridge regression.

aeENR: elastic net regression.

afDNN: deep neural network.

anNP: neuropathic pain.

aLS: least squares.

a3D-RCNN: 3D region-based convolutional neural network.

aC-CCA: nonlinear canonical correlation analysis.

aLRL1: L1 penalized logistic regression.

aLRL2: L2 penalized logistic regression.

aLR-EN: logistic regression-elastic net.

aET: extra tree.

aGTB: gradient tree boosting.

aML: machine learning.

aELM: extreme linear machine.
average sample size was 1686, and the studies varied in sample size: <100 participants (8/42, 19%), between 100 and 1000 participants (27/42, 64%), and >1000 participants (7/42, 17%). Most studies relied on clinical data (28/42, 67%), although some integrated clinical data with computed tomography (CT) images (14/42, 33%). The study designs were diverse, including retrospective (18/42, 43%), cross-sectional (15/42, 38%), prospective (5/42, 12%), and longitudinal (4/42, 10%) approaches.

Cancer Primary Sites and Predicted Symptoms

Various primary cancer sites were studied, with head and neck cancers being the most prevalent (9/42, 21%). Breast cancer was the focus of 19% (8/42) of the studies, and lung cancer was studied in 17% (3/42) of the cases. The included studies included participants undergoing a range of treatments, including chemotherapy (9/42, 21%), radiotherapy (9/42, 21%), surgery (4/42, 10%), and investigations of posttreatment survivors (2/42, 5%). Of the 42 included studies, 10 unique symptoms were reported as outcome variables in the predictions. Those included were xerostomia (9/42, 14%) [27,30,34,49-51,55,58,60], depression (8/42, 13%) [22,33,37,41,43,45,52,59], pain (8/42, 13%) [20,25,26,35,37,40,42,56], fatigue (6/42, 10%) [23,24,29,38,42,44], anxiety (3/42, 5%) [33,46,52], sleep disturbance or insomnia (3/42, 5%) [26,28,52], nausea or vomiting (3/42, 5%) [17,26,29], weight loss (2/42, 3%) [47,53], cognitive impairment (2/42, 3%) [21,36], and diarrhea (2/42, 3%) [29,42].

One study reported multiple symptoms, including hypersensitivity [29], stomatitis [29], hand-foot syndrome [29], peripheral neuropathy [29], and constipation [29]. Another study delved into taste and general activity [40]. Individual studies were dedicated to each of the following symptoms: delirium [31], lung infection [32], lymphedema [39], well-being [37], odynophagia [54], social distress [26], spiritual pain [26], dyspnea [26], and hearing loss [57]. The distribution of these symptoms is depicted in Multimedia Appendix 5.

Significant Candidate Predictors of Symptoms

Numerous predictors were frequently used for predicting symptoms, which can be grouped into demographic features and clinical characteristics.

Demographic Features

The demographic features include age, sex, BMI, income, medical insurance, education, marital status, and zip code–level poverty.

Clinical Characteristics

The clinical characteristics include smoking and alcohol use, initial diagnosis, presence of cancer, stage of cancer, cancer course, tumor site, type and number of prior treatments, chemotherapy type, and radiotherapy dose and volume. Health conditions such as comorbidity, diabetes, hypertension, osteoarthritis, and coronary disease also play a significant role. In addition, psychological factors such as depression and anxiety, fatigue, sleep disturbance, and pain are considered. Other influential predictors encompass care fragmentation, polypharmacy, hormone levels, physical activity, diet, heart rate, and social support factors.

In our comprehensive analysis of 42 studies, all the detailed findings on common cancer symptoms are compiled in Figure 2. We provide a detailed analysis of the predictors for the 4 most frequently reported cancer symptoms identified in this study: xerostomia, pain, depression, and fatigue. In a detailed analysis of 42 studies, various predictors for 4 common cancer symptoms—xerostomia, pain, depression, and fatigue—have been identified, each with its distinct set of influencing factors.
For xerostomia, age, gender, chemotherapy type, radiotherapy dose and volume, cancer site, tumor size, and hypertension are crucial predictors. In the case of pain, factors such as age, BMI, smoking and alcohol habits, cancer site and stage, tumor size, diabetes, hypertension, osteoarthritis, coronary disease, physical activity, psychological factors, sleep disorders, and existing pain conditions emerge as significant. Significant predictors for depression include age; gender; education; cancer site and stage; economic factors such as insurance, income, and poverty level; marital status; initial diagnosis impact; comorbidities (diabetes, hypertension, osteoarthritis, and coronary disease); pain; social support; care fragmentation; polypharmacy; and various scale scores. Finally, for fatigue, the key predictors are existing fatigue and low energy, cancer site, sleep disturbances, age, income, education, chemotherapy type, tumor size, comorbidities, hypercholesterolemia, heart rate, hyponatremia, physical and psychological factors, pain, adverse drug reaction history, limited social support, Eastern Cooperative Oncology Group score, platelet distribution width, and erythropoiesis.

When examining the commonalities across these predictors for xerostomia, pain, depression, and fatigue, several factors stand out as particularly influential across multiple symptoms: age; gender; cancer site and stage; treatment-related factors such as the type of chemotherapy and radiotherapy; comorbidities such as diabetes, hypertension, and coronary disease; physical and psychological factors; and socioeconomic factors such as income and education level, demonstrating the impact of cancer treatments on symptom development. These common predictors underscore the complex, multifactorial nature of symptom manifestation in patients with cancer, necessitating a comprehensive approach to their management and care.

**ML Algorithms and Validation Metrics**

Of the 42 studies analyzed, 7 (17%) used a single ML algorithm, whereas 35 (83%) used multiple algorithms. The most effective models, in terms of performance, were logistic regression (LR; 9/42, 17%), random forest (7/42, 13%), artificial neural networks (5/42, 9%), decision trees (DTs; 5/42, 9%), and extreme gradient boosting (3/42, 6%). For validation methods, 10-fold cross-validation was the most used (14/42, 31%), followed by 5-fold cross-validation (5/42, 11%), 3-fold cross-validation (2/42, 4%), and 8-fold cross-validation (1/42, 2%). The primary evaluation metric across these studies was the area under the curve, which was adopted in 24% (26/42) of the studies. A visual representation of the leading ML models along with the validation and evaluation metrics used in the study presents in Multimedia Appendix 6.

**Discussion**

**Principal Findings**

In this review, we present the first systematic analysis of ML applications for predicting the development of cancer symptoms. We explore the most frequently studied cancer sites and delve into the intricacies of ML procedures. Breast, head or neck, and lung cancers are the most frequently studied sites in current research, with xerostomia, depression, pain, and fatigue being the most prominent symptoms. The application of various ML techniques is on the rise, with data acquisition and preprocessing being pivotal for successful ML models. While a range of
algorithms, from traditional methods such as LR and DT to advanced ones such as DL, are used, there is a growing emphasis on data quality, external validation, and a standardized approach to model evaluation. The future of ML in cancer symptom prediction looks promising, with a need for collaborative efforts among oncologists, data scientists, and patient groups, combined with more comprehensive research on lesser-studied cancer sites and standardized methodologies.

Regarding the cancer sites covered in the studies, breast, head, and neck cancers emerged as the most frequently researched primary cancer sites. The range of symptoms and side effects that patients experienced varied from one study to another. Some symptoms depended on the specific cancer site and the treatments patients received. For example, xerostomia, which can either arise from the tumor itself or manifest as a treatment side effect, has a significant impact on patients’ dental health and compromises antimicrobial functions [61]. However, most symptoms were not directly attributed to a particular cancer site or treatment.

Our review revealed a notable emphasis on predicting xerostomia in 14% (9/62) of the studies, despite head and neck cancers being less prevalent. The notable emphasis on predicting xerostomia in ML research, despite the lower prevalence of head and neck cancers, is due to advancements in integrating ML with CT imaging. CT imaging is a pivotal tool in the diagnosis and treatment planning of head and neck cancers. The integration of ML with CT imaging has opened new possibilities for more accurately predicting side effects such as xerostomia. ML techniques, when applied to CT images, can potentially identify patterns and indicators that are not easily discernible by human observers. This capability can lead to earlier and more precise predictions of xerostomia, thereby enabling better preventive measures and treatment planning to mitigate this side effect. Therefore, the focus on xerostomia in ML research, in the context of head and neck cancers, is likely driven by the opportunities presented by combining ML with advanced imaging techniques.

Depression, a widespread emotional challenge for people with cancer [62,63], was the focus of prediction in many studies (8/24, 13%). Similarly, pain, a recurrent concern for palliative care patients [64] and survivors of cancer [65,66], was the subject of prediction in >13% (8/24) of the studies. Fatigue, prevalent across all age groups with cancer [67,68], was highlighted in 6 (10%) of the 42 studies reviewed.

In terms of the ML approaches used in the studies, a plethora of techniques were used to construct these predictive models, spanning all phases of the ML process, from data collection and preprocessing to feature and algorithm selection, model training, testing, and evaluation. The process of data acquisition is pivotal for the development of ML models, thereby emphasizing the importance of an adequate sample size. Upon reviewing 42 studies, we discerned that the most frequent sample sizes for ML applications ranged between 100 and 1000 samples. More advanced ML techniques necessitate larger data sets to bolster robustness and mitigate the risk of overfitting. Alarmingly, certain studies in our review used ML with comparably smaller data sets, introducing the risk of model overfitting and potential biases in the subsequent performance metrics [69]. Challenges tied to sample size might impede the creation of sturdy and trustworthy ML models [70]. Data preprocessing is indispensable to yield clean and interpretable data, which is a cornerstone for proficient ML models. Data cleaning approaches encompass addressing missing values, tackling data noise, and data normalization. Within health care data sets, noisy or absent data are frequently a by-product of inaccuracies in manual entries or instrument recordings made by medical personnel or ancillary staff [71]. However, most of the reviewed studies lacked comprehensive descriptions of their data cleaning methodologies or strategies for handling noisy data and normalization, constrained by word or page limits in publications.

Given the crucial importance of data quality in developing ML models, it is essential for researchers to focus equally on effective data preparation and choosing suitable algorithms. Future endeavors would benefit from exhaustive procedural documentation made available on public platforms such as GitHub. In a research context, GitHub can be used for sharing and collaborating on various aspects of a research project, including but not limited to code. It allows researchers to maintain version control of their scripts, data analysis procedures, and even documentation. This feature is particularly beneficial for replicating studies and verifying results, as it provides a transparent view of the methodologies and analyses used.

Overloading an ML model with excessive features can undermine its ability to differentiate between pertinent data and superfluous noise, leading to the challenge often referred to as the “curse of dimensionality.” The goal of feature engineering is to mitigate model complexity, expedite the training process, reduce the data’s dimensionality, and avert overfitting [72]. By streamlining the model with a curated set of predictors, it becomes more accessible and transparent, emphasizing the importance of feature selection during data preparation. Our review pinpointed the most frequently used significant predictors in cancer symptom prediction. The efficacy of prediction models is heavily influenced by the number and interplay of the relevant predictors. Factors such as age, gender, type and number of previous treatments, cancer location, cancer stage, chemotherapy type, dosage and volume of radiotherapy; chronic conditions such as diabetes and hypertension; concurrent diseases; and symptoms including depression, anxiety, fatigue, pain, and sleep disturbances have consistently featured as determinants in numerous predictive frameworks. Our review of cancer symptom prediction underscored age as a pivotal factor, associated with predominant symptoms such as depression, pain, xerostomia, and fatigue. While numerous elements, from gender to type of treatment and cancer stage, influence the predictive models, it is the prominence of age that consistently emerges as a cornerstone predictor. As we delve deeper into this field, even with the introduction of newer determinants and correlations, the centrality of age in these frameworks remains indisputable.

Regarding algorithm selection, traditional methods often struggle with handling high-dimensional data and processing extensive information. To tackle these challenges, researchers have
increasingly shifted toward innovative ML algorithms that are renowned for their robust predictive power and strong generalization capacities. These sophisticated algorithms excel at delving deep into data and discerning intricate interrelationships among variables. To navigate the multifaceted landscape of modeling challenges, it is advantageous for researchers to leverage a diverse array of ML algorithms. Most studies used multiple predictive models, with techniques such as LR, RF, ANN, and DT consistently delivering stellar results. The introduction of advanced ML techniques, such as DL and ensemble classifiers, provides promising opportunities to elevate prediction accuracy in future research.

After their design, the ML models undergo training and testing on different data sets. However, these models can grapple with issues such as overfitting and underfitting. Overfitting occurs when a model becomes overly complex, which leads to increased variance and reduced clarity. In contrast, underfitting results from an oversimplified model, causing it to overlook key data patterns and diminish its predictive capacity. Therefore, the ideal learning model should strike a balance between the optimal variance and justifiable bias. To mitigate these issues, the common strategy is to divide the data set into training and testing subsets, followed by internal or external validation. While most studies in our review used internal validation, only 1 study reported external validation [58], which was demonstrated on a small cohort of 25 patients with head and neck cancer. Although its performance is typically lower than evaluations using the original data sets, external validation remains crucial for gauging ML models [72]. It is a crucial step in ensuring that the model’s performance is not just limited to the conditions and data it was originally trained on but also applicable and reliable in broader, real-world clinical settings. This approach serves to verify the model’s efficacy and generalizability across different patient populations and settings.

Understanding and interpreting ML models continue to pose challenges. Determining the variables that significantly impact symptom prediction can be elusive due to the intricate prediction processes. Many studies gauge the performance of ML models using metrics that examine their ability to distinguish between 2 classes. From our systematic review of 42 studies, the area under the curve emerged as the predominant metric for the prediction models. Other metrics included accuracy, sensitivity, specificity, positive predictive value, root mean square error, and negative predictive value. These metrics provide a holistic view of a model’s efficacy, facilitating its refinement and enabling more precise predictions. However, the diverse emphasis on distinct metrics in numerous studies underscores the need for a uniform approach to evaluating ML models in cancer symptom prediction.

As interest grows in using ML for predicting cancer symptoms, there are several areas that merit deeper investigation. A crucial area is broadening the range of studied cancer sites and more comprehensively correlating symptoms with various treatment methods. To fully understand symptom prediction, it is essential that future studies delve into lesser-explored or infrequently studied cancer sites. Furthermore, the methodologies used for data preprocessing and cleaning should be documented more thoroughly, focusing on best practices to ensure data integrity. As data are foundational to ML models, transparent and detailed preprocessing can improve the reliability and repeatability of these models. Although our analysis highlighted common predictors for symptom forecasting, examining potentially underrepresented or emerging indicators could refine these models further. On the algorithmic front, exploring hybrid ML methods that merge the strengths of multiple algorithms might be particularly beneficial for cancer symptom prediction. Standardizing evaluation metrics across studies would also provide clarity and facilitate a more accurate comparison of various ML techniques. To genuinely progress, collaborations among oncologists, data scientists, and patient advocacy groups are vital to ensure that the developed models are technically robust and clinically pertinent. With these insights, ML stands poised to transform cancer care, creating treatment plans based on patient-focused and accurate symptom prediction models.

Limitations
This review is not without its limitations. Although we established clear inclusion and exclusion criteria, potential biases in the studies we analyzed could inherently limit our review. We might have missed or excluded relevant studies due to inadequate information or the absence of keywords in their titles or abstracts. Many of the studies we reviewed did not specify the cancer site, potentially limiting the accuracy and applicability of our findings to specific cancer types. The broad range of predictors used across the studies also made it difficult to draw definitive conclusions about the most influential factors in predicting cancer symptoms using ML algorithms. As such, readers should interpret these results cautiously, given this variability.

Conclusions
ML offers an intriguing potential for predicting cancer symptoms, thereby preemptively mitigating the associated challenges. Predicting the symptoms that people with cancer might experience and determining their onset throughout their treatment journey is a pivotal clinical issue that can enhance patients’ quality of life. Notably, all studies in our review were published after 2017, highlighting the nascent nature of this research area. Our investigation primarily sought to outline the ML methodologies harnessed for symptom prediction in people with cancer. While ML techniques hold an edge over traditional statistical approaches by virtue of their prowess in analyzing vast data sets and gauging the efficacy of diverse prediction models, certain impediments such as a limited pool of symptoms; suboptimal data preparation; challenges in feature engineering; and complexities in ML algorithm design, validation, and evaluation can constrain the broad applicability of these predictive models. Future research should pivot toward amplifying the efficacy of ML strategies. This enhancement can be achieved by harnessing expansive, high-caliber data sets; tapping into innovative technologies for data refinement; and sculpting refined models. Harnessing ML can potentially free health care practitioners—including doctors, nurses, and clinic personnel—to accentuate the human touch in managing cancer symptoms.
Acknowledgments
The authors would like to express their profound gratitude to their esteemed colleagues and academic mentors. Their invaluable insights, unwavering support, and dedicated time significantly enriched the authors' research journey. Special acknowledgment is reserved for the myriad of researchers and study participants whose dedication and data have underpinned this comprehensive review. Their collective efforts have made this work possible. The authors extend special thanks to Jennifer Deberg, a specialist at the Hardin Library for the Health Sciences, for her invaluable support in selecting search terms and databases.

Conflicts of Interest
None declared.

Multimedia Appendix 1
The detailed search strategy for the databases and the Boolean expressions used.
[DOCX File, 30 KB - cancer_v10i1e52322_app1.docx ]

Multimedia Appendix 2
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.
[PDF File (Adobe PDF File), 942 KB - cancer_v10i1e52322_app2.pdf ]

Multimedia Appendix 3
The distribution and overlap of 42 studies across the databases.
[PNG File, 88 KB - cancer_v10i1e52322_app3.png ]

Multimedia Appendix 4
The data extracted from 42 studies.
[DOCX File, 31 KB - cancer_v10i1e52322_app4.docx ]

Multimedia Appendix 5
Number of studies per cancer symptoms.
[PNG File, 62 KB - cancer_v10i1e52322_app5.png ]

Multimedia Appendix 6
Visual overview of the machine learning models and metrics.
[PNG File, 74 KB - cancer_v10i1e52322_app6.png ]

References


Abbreviations

CT: computed tomography
DL: deep learning
DT: decision tree
LR: logistic regression
ML: machine learning
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

©Nahid Zeinali, Nayung Youn, Alaa Albashayreh, Weiguo Fan, Stéphanie Gilbertson White. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Viewpoint

Need for Culturally Competent and Responsive Cancer Education for African Immigrant Families and Youth Living in the United States

Olufunmilola Abraham¹, BPharm, MS, PhD; Adeola Agoke², MA, PhD; Kazeem Sanuth³, MA, PhD; Abimbola Fapohunda⁴, MPH, MS, DrPH; Motolani Ogunsanya⁵, BPharm, PhD; Megan Piper⁶, BA, MClinPsych, PhD; Amy Trentham-Dietz⁷, MS, PhD

¹Social and Administrative Sciences Division, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, United States
²African Cultural Studies, University of Wisconsin-Madison, Madison, WI, United States
³National African Language Resource Center, Indiana University Bloomington, Bloomington, IN, United States
⁴Behavioral and Community Health Sciences, School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States
⁵College of Pharmacy, Health Sciences Center, University of Oklahoma, Oklahoma City, OK, United States
⁶Department of Medicine and Center for Tobacco Research and Intervention, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, United States
⁷Population Health Sciences and Carbone Cancer Center, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Olufunmilola Abraham, BPharm, MS, PhD
Social and Administrative Sciences Division
School of Pharmacy
University of Wisconsin-Madison
Room 2515 Rennebohm Hall
777 Highland Avenue
Madison, WI, 53705
United States
Phone: 1 6082634498
Fax: 1 6082625262
Email: olufunmilola.abraham@wisc.edu

Abstract

Cancer prevalence data for Black Americans is monolithic and fails to consider the diverse cultures and backgrounds within that community. For instance, African immigrants constitute a meaningful proportion of the foreign-born Black immigrants in the United States (42%), but the prevalence of cancer in the African immigrant community itself is unknown. Therefore, without accurate cancer prevalence data, it is impossible to identify trends and other key factors that are needed to support the health of African immigrants and their children. Moreover, it is impossible to understand how the culture and language of subgroups influence their cancer-related health behavior. While research in this area is limited, the existing literature articulates the need for culturally responsive and culturally tailored cancer education for African immigrants and their adolescent children, which is what we advocate for in this viewpoint paper. Existing projects demonstrate the feasibility of culturally responsive programming for adults; however, few projects include or focus on adolescents or children born to African immigrants. To best meet the needs of this understudied community, researchers must use culturally competent interventions alongside familiar, usable media. For adolescents, technology is ubiquitous thus, the creation of a culturally tailored digital intervention has immense potential to improve cancer awareness and prevention for youth and their community. More research is needed to address many of the existing research gaps and develop a rich understanding of the unique experience of cancer among African immigrant families that can be used to inform intervention development. Through this viewpoint, we review the current state of cancer-related research among African immigrant families in the United States. In this paper, we acknowledge the current knowledge gaps and issues surrounding measurement and then discuss the factors relevant to designing an educational intervention targeted at African immigrants and the role of African immigrant youth.

(JMIR Cancer 2024;10:e53956) doi:10.2196/53956

https://cancer.jmir.org/2024/1/e53956
KEYWORDS
African immigrant; youth; adolescent; adolescents; teen; teens; teenager; teenagers; cancer awareness; health disparities; culturally competent education; African; Black; immigrant; cultural; culturally; cancer; oncology; patient education; awareness; prevention; disparity; disparities

Introduction
Cancer has a profound impact on the experience of health for many in the United States that only continues to grow. Research has demonstrated the escalating rates of early-onset cancer diagnosis among women and the alarming decreasing rates for men and Black people; most commonly in breast, thyroid, and colorectal cancer [1]. Early diagnosis and prompt treatment of cancer are critical to improved public health. The decreasing rate of early cancer detection and response is introducing a significant health inequity among Black people in the United States. African-born immigrants and their children comprise a meaningful portion of the US population. The paper aims to describe the existing research gaps and experiences of cancer among African immigrant families and highlight the need to design and tailor cancer education for African immigrant families.

There was a surge in the African immigrant population between 1970 and 2015 [2]. This migration pattern has continued, with the African immigrant population growing from 881,000 in 2000 to 2.0 million in 2019, comprising 42% of the US foreign-born Black population. African immigrants have tended to settle in 4 main cities in the United States: Washington DC, New York City, Minneapolis or St Paul, and Atlanta [3]. Prior research has established that most African immigrants come from Western (35.7%) and Eastern Africa (29.6%) [4]: from Nigeria (13.7%), Ethiopia (10.8%), Ghana (7.8%), and Kenya (5.5%) [5]. Therefore, Nigeria and Ethiopia constitute the top birthplaces of African immigrants in the United States.

In 2011, one of the first peer-reviewed papers on African immigrant health was published. It highlighted the growing population of African immigrants in the United States and the need to study their health care needs and practices since chronic diseases, including cancer, were poorly understood [3]. There is a growing research interest in African immigrant health, especially by researchers who are first- and second-generation African immigrants themselves, especially in light of the myriad of factors that impact African immigrants’ health, and that of their families, including the trauma of immigration, conflicting cultural contexts between African family dynamics and those common in the United States, diet and lifestyles, culture, religion, and spirituality. These constructs make up the richness of immigrants’ lives and continue to impact their health behaviors, health care experiences, and decision-making about their health practices after they move to the United States. Therefore, attention must be given to these factors. These factors also influence how African immigrants interact with and navigate the health care system, making it essential to understand how these factors can negatively impact health care system involvement.

The number of health-related areas influenced by immigration only grows as their length of stay in the United States increases [6]. Given the unique health experiences of African immigrants, the need to better understand the cancer-related health of African immigrants is imperative. The United States has begun to address disparities in immigrant health, such as affordances of health care following the implementation of the Affordable Care Act or state-level legislation allowing immigrants (especially the young, pregnant persons, and older people) to be eligible for state Medicare. States that have expanded care for immigrant children have seen reductions in no-insurance rates and rates of families forgoing medical care [7]. However, there is a current dearth of initiatives directly designed for African immigrants. With a deeper understanding of cancer in the African immigrant community, we can create novel, innovative, and culturally specific educational tools to support African immigrant families and improve current and future African immigrant community cancer health.

Discerning Cancer Prevalence Among African Immigrants

Overview
Uncovering cancer prevalence and awareness among African immigrants is challenging. Racial and ethnic minority groups are underrepresented in health research overall, contributing to persistent health disparities in the United States [8]. Cancer death rates among Black people continue to increase and so does the risk of developing cancers [9]. In the past few decades, there have been focused efforts to conduct research and draft policies to address health disparities within immigrant communities. However, there is a discernable lack of research on African immigrants’ (and their children’s) health related to cancer prevention and awareness in the United States. The challenge is due to limited resources allocated to minority issues and a lack of effort to distinguish the population as different and unique from other minority populations [10,11].

Issues With Measurement
Most research on immigrant health in the United States has focused on Latinos and Asian Americans [12-15]. Similarly, most cancer-related research in the United States has used race, and Hispanic or Latino ethnic affinity, regardless of the differing histories of migration, as the basis of categorizing research participants. Therefore, there is limited knowledge about African immigrant health in the United States, especially on cancer awareness, cancer care, and overall health outcomes. Accurate prevalence of cancer among African immigrants is unknown, available literature mostly focuses on databases that have combined data for African-born immigrants and US-born Black people. This makes it difficult to identify African immigrants and to provide accurate evidence of the extent and impact of cancer within their communities [16]. This practice facilitates a monolithic view of people with African heritage; therefore, it discourages granularity of analysis and limits health services’
researchers’ ability to address African immigrant-specific health challenges and examine related research questions.

Some researchers have started to address the overgeneralization of categorizing all African immigrants as “Black” by focusing on their country of origin or time since immigration or assimilation or acculturation [17,18]. Assimilation allows immigrants to integrate into the social, linguistic, and cultural fabric of the host society. However, acculturation experience differs across immigrant groups. The Hispanics, specifically Mexican Americans, constitute an immigrant group in the United States with a robust acculturation. Safran Williams in his classification of diasporas describes Mexican American as “not true diaspora” [19]. This is because of their immigration history with the United States [19]. Further, Spanish is the dominant language of the Mexicans and is also the most popular foreign language in the United States. As a result, acculturation for Mexican immigrants is steady and impacts the strength of research and health interventions for this immigrant group [20-22]. Contrarily, the cultural and linguistic significance of African immigrant identities do not share the same history and recognition both in the United States social milieu and in the US health care system especially those relating to cancer education and research. African immigrants have an existing cultural identification from their homeland and their languages do not have the same recognition as that of Hispanic Americans. Nonetheless, the effort to acculturate among African immigrants accounts for the experiences such as changes in diet, modified language practices, and using the health care services for access to information, treatment, and care. The acculturation process is also layered with the African immigrants’ spirituality and how it influences their reception of health care treatment. Careful attention to the cultural practices of African immigrants and their relevance to health intervention will largely impact the outcomes in cancer awareness and education.

The issue of having a monolithic “Black” category affects the extraction of research data on African immigrants [16]. Some progress is being made in this area. For instance, I study promoted awareness and accessibility to screening for chronic diseases among African immigrants living in Georgia [17]. Other research has discussed African immigrants’ health and allostatic load score as it relates to cardiovascular, metabolic, and immune systems [18]. Finally, a scoping review identified additional socio-ecological challenges faced including the lack of culturally competent health care, distrust of the health care system, challenges navigating the US health system, and the burdensome cost of care [16].

What We Do Know About African Immigrants and Cancer

Accurate prevalence rates of cancer in African immigrants in the United States are lacking. Evidence suggests high cancer prevalence in their countries of origin, especially breast and cervical cancer for women and prostate cancer for men [23-25]. More research is needed to understand the prevalence of cancer within immigrant families and how their immigration may influence cancer prevalence.

The experiences and needs of African immigrants are unique [17,26-32]. Sociocultural factors underlie the experience of cancer in the African immigrant community. The stigma of being diagnosed with cancer, lack of cancer awareness, limited or no screening (especially among African immigrant women), and limited familiarity with prevention strategies and treatment technologies available may be contributing to the high prevalence of cancer [24,30]. These factors lead to late-stage diagnoses because of a lack of access to health care, lower education levels, and cultural and religious beliefs regarding cancer [33,34]. Studies also found that African-born women have limited knowledge and exposure to breast cancer screening information before their arrival in the United States [30,34,35], which can impact their preventative and cancer screening behaviors. Existing research has also explored cancer mortality among adults across different Black ethnic groups—African, African American, and Caribbean—showing some mortality and prevalence differences between these groups [36].

Further, I study has found that income, among other factors, plays a significant role in the population’s understanding of colorectal cancer [37]. With a focus on breast and cervical cancer screening, other studies examined the knowledge and perspectives of African immigrants [38,39]. Their findings underscore significant factors impacting the decision to seek preventative screening measures among African immigrants, including fatalism, lack of cancer knowledge, stigma, length of stay in the United States, provider gender, and privacy concerns [40-43]. Another study examined prostate cancer risk experiences among West African men and shed light on the modifiable risk factors implicated in prostate cancer mortality and morbidity [44]. A study of cervical cancer awareness among African immigrant women in Iowa City highlighted factors such as fear, languages spoken, and education as barriers to preventative treatment [45]. Considering the available research and prevalent factors that limit cancer prevention knowledge and behavior it is imperative to develop culturally, and linguistically appropriate cancer education programs aimed at increasing awareness and screening of cancer. In summary, while research has begun to address differences in African immigrant health, the differences are many which will require further study and consensus.

Lack of Cancer Awareness Among Youth and African Immigrants

In 2008, it was estimated that the 82% of the US population increase between 2005 and 2050 would be attributed to immigrants and their descendants [26]. Despite an increase of African immigrants’ offspring in the United States, little is known about these second-generation individuals born and raised in the United States (with at least 1 foreign-born parent), regarding their health beliefs, perceptions, and practices. This is understandable as little is known about their parents regarding these factors. A study that explored beliefs and lifestyle behaviors relating to healthy living and diet among middle-aged adults in the immigrant population indicated that little is known about the beliefs, perceptions, and practices of diet and exercise among young African immigrants [46]. Young adults of African
immigrant descent are part of the future, and attention needs to be paid to their well-being.

It is unknown if children of African immigrants are being educated about cancer by their parents, their communities, their health care providers, or in schools. Cancer is often termed as a taboo subject in most African homes and communities. This is further compounded by other barriers such as access to care, quality of care, communication gaps, lack of education, lack of affordable health care, lack of transportation, socioeconomic status, shame and stigma, and cultural and religious beliefs [47]. Nonetheless, some children of African immigrants become aware of cancer when close family members or friends are diagnosed. With limited cancer awareness and the vulnerability of African immigrants regarding cancer, youth, and their parents must be educated using culturally competent, tailored, and responsive family-oriented cancer education initiatives that build on the strengths of these immigrant cultures as well as address the barriers to cancer prevention behaviors.

Although the limited research reviewed above examines cancer among Black immigrant men and women, there is no substantial body of research that addresses cancer education and awareness among first and second-generation African immigrant adolescents in the United States. A lack of knowledge about youth immigrants and second-generation African immigrants can put this population at a disadvantage as compared to their peers. Cancer awareness among African immigrants and youth studies, including older and younger Somali women, use age as a factor for examining standardized prevalence of cardiovascular disease risk factors among both African immigrants and African Americans [47,48]. Although age is an important factor to consider, this work does not focus on youth. Another study, rather than age, used the year of residence in the United States to examine self-reported health problems among African immigrant adults [49].

While several studies have begun to address cancer research among the African population broadly, the significant paucity of research that focuses on the youth of African immigrant families in the United States leaves a critical gap in cancer awareness and prevalence research. To our knowledge, no studies have sought to examine or address cancer awareness among the youth of African immigrant families, nor interventions for cancer awareness and education. The youth of African immigrant families in the United States constitute an important population that is instrumental in creating awareness about the prevalence of cancer within their community. To access the youth groups of African immigrant descent in the United States, it is expedient to identify cultural and age-relevant educational tools for creating awareness about the prevalence of cancer disease.

Existing Studies on the Promotion of Cancer Awareness and Education Among African Immigrants

Overview

There is evidence of studies that promote cancer health education among African immigrants and other minority groups using various culturally tailored approaches and technologies. The success of a community-academic partnership model at community faith-based centers is effective for immigrant women in learning about breast cancer [50]. Moreover, health education programs in community-based settings have indicated strong potential. Further, 2 studies involving interpreters and culturally targeted communication, showed increased breast cancer knowledge and an improvement in screening for breast cancer for immigrant and multicultural women [51,52].

Study findings have demonstrated the importance of culturally tailored educational tools and different approaches to reduce cancer-related disparities. These studies provide strong evidence supporting the use of culturally relevant educational materials, patient navigation programs, peer-to-peer education, education programs, videos, and cofacilitated health promotion forums in promoting preventative and cancer screening behaviors [33,53-62]. Together these projects shed light on some of the few, yet variable opportunities for successful community-engaged research with African immigrant families.

Furthermore, some studies have demonstrated the potential of technology in promoting cancer awareness and education among African immigrants. Mobile devices, tablets, and computers have been used to address common cultural and linguistic barriers to cancer screening. Mobile health initiatives, culturally tailored messaging, language support, mobile apps, short message services, and text messages have all proven effective in impacting cancer screening behaviors [18,63-65]. Some of these initiatives could be adapted into family-based programs where young African immigrants could learn in familiar spaces using ubiquitous and widely acceptable technologies such as serious games.

Global Health Perspectives and Solutions for Culturally Competent Care Among African Immigrants

Health care approaches for immigrant populations require adaptation and cultural competence to serve diverse communities effectively. Parallel analysis of the US health care models with those of other nations like Canada and Australia offers a framework to evaluate and refine strategies to address health disparities among African immigrants. Canada and Australia have made strides in fostering inclusive health strategies that can inform US health care practices, particularly in providing culturally competent care to African immigrants.

For example, in Canada, health care delivery to immigrant populations acknowledges the necessity of cultural competence. Canada’s universal health care system actively integrates culturally tailored interventions. The Canadian government has pushed for strategies that involve community engagement and representation in health decision-making, enhancing the cultural appropriateness of health care services [66]. Using community health workers who share the same cultural background as immigrants has been a breakthrough, acting as a bridge between health care providers and immigrant communities [67]. These community health workers facilitate communication, understanding, and trust—essential elements in promoting the health and well-being of immigrant populations [68].
Further, Australia’s approach to immigrant health pivots on inclusivity and health equity to deliver services that are respectful of and responsive to diverse patients’ health beliefs, practices, and needs [69]. A notable instance is the Victorian Immigrant and Refugee Women’s Coalition’s efforts, which engage women directly to educate about health issues, including cancer awareness [70]. Australian health policies aim to address the language barriers and the diverse cultural contexts that can influence health care usage and outcomes. In contrast, the United States continues to grapple with creating a standardized approach for culturally competent care throughout its health care system.

While there are pockets of exemplary practices, such as using patient navigators in cancer care to assist patients from minority backgrounds, there is not a universal health care mandate specifically aimed at immigrant health [67]. Instead, the United States relies on a patchwork of local initiatives and federal guidelines, such as those by the Office of Minority Health which established the National Standards for Culturally and Linguistically Appropriate Services in health and health care [71]. In conclusion, both the Canadian and Australian models underscore the importance of cultural competence and systemic support in improving immigrant health outcomes. They demonstrate that effective immigrant health strategies require the integration of culturally informed practices across all stages of health care—from preventive education to treatment. This implies adopting multifaceted approaches that can cater to the unique cultural, linguistic, and religious elements that define African immigrant communities.

Youth: the Bridge for Culturally Tailored Cancer Education

Given their positionality, first through second-generation African immigrant youth are at a unique nexus from which they can bridge health gaps related to cancer that arise from their heritage and sociocultural contexts. Cultural tailoring acknowledges the broad culture but identifies specific strategies for reaching specific individuals. These groups of individuals have insights into the linguistic and cultural practices of their families as well as those of the society they live in. Due to their positionality, the youth are motivated to embrace language awareness, which emphasizes the interrelatedness of language, culture, and social structures [72]. The interrelatedness of cultural meanings and linguistic signs allows for the tailoring of educational content that addresses distinctive groups. The adolescents of African immigrant families are a product of the diverse linguistic and cultural interactions that occur through transnational migration and globalization.

To engage with youth and form a robust bridge between coexisting sociocultural systems to improve African immigrant community health, research should focus on methods that are familiar and usable for adolescents. A ubiquitous facet of adolescent life is technology. There is increasing interest in serious games (i.e., games that serve an educational or developmental purpose aside from pure entertainment) as a learning medium. Although innovative interventions including serious games are becoming popular, they are not traditionally designed and tailored to meet the cultural and health needs of minoritized populations such as African immigrant families. Systemic reviews of serious games indicate limitations that need to be addressed [73-76]. It will be beneficial for health services’ researchers to use a participatory design approach when designing cancer education and intervention tools for African immigrant families. Such a collaborative approach will allow African immigrant families to partner in the co-design of technologies such as serious games and facilitate the creation of a culturally competent and responsive learning medium. Youth from African immigrant families typically have a hybrid of identities which necessitate the use of education technologies such as serious games in ways that speak to their lived experiences and families’ cultural heritage and realities. Therefore, there is a need to tailor educational resources using technology platforms that would engage the linguistic and sociocultural realities of the African immigrant population. Interventions to improve cancer outcomes in African immigrants, especially among youth, are necessary.

Youth and community members from other minority populations in the Northwest Arctic region of Alaska participated in community-based participatory action research honoring indigenous ways, creating a Sharing Circle used to understand community priorities and develop culturally relevant cancer education that could be incorporated into school curriculum. It is an opportunity for youth involvement in culturally relevant health promotion efforts to address health disparities in cancer [77].

Culturally Tailored Education for African Immigrant Youth

Overview

Developing educational resources for African immigrant youth brings into focus the question of curricular content and pedagogical approaches that fit this group. The connection of educational content with cultural identities is espoused in the framework of culturally relevant pedagogy (CRP) [78]. CRP encompasses multiple concepts related to students’ academic achievements and social inequalities, but its central tenet is the interconnection of theories and cultures in manners that will “empower students intellectually, socially, emotionally, and politically by using cultural referents to impart knowledge, skills, and attitudes [79].” African immigrants and people of historically marginalized cultures are unique and deserving of an educational approach that is aligned with their needs. It offers liberatory education which inspires the learners to become social commentators, advocates, and critical consumers of knowledge while empowering control over one’s health. The use of such an approach will be beneficial in disseminating and promoting cancer education in the community.

The pedagogical approaches to achieving culturally tailored education may derive from CRP and adopt effective strategies that will merge critical consciousness and cultural connections in the learning content. CRP proposes three components that must be integrated to achieve learning: (1) a focus on youth learning and academic success, (2) developing youth’s cultural
competence to assist them in developing positive ethnic and social identities, and (3) supporting youth’s critical consciousness or their ability to recognize and critique societal inequalities.

Researchers have described examples of targeted and tailored strategies, techniques, and procedures for successful intervention with a variety of populations [80]. These researchers identified linguistic, community-engaged, and sociocultural strategies as important to reaching a particular community. Building on this knowledge, we identify four approaches that a cancer education intervention that the youth of African immigrant heritage can draw on, namely: (1) linguistic and cultural markers, (2) belief system and religious affiliation, (3) hybrid nationality, and (4) age-related learning preferences. With a deeper understanding of how these factors, concerning cancer health, shape the identities, beliefs, and behaviors of African immigrant youth in the United States, we may be able to create culturally competent educational tools for cancer awareness and prevention.

**Linguistic and Cultural Markers**

African immigrants, having come from different countries with diverse colonial histories, have distinct languages. The native languages of African immigrants play an important part in their identity. The youth of African immigrants assimilate the language and cultures of the host society while leveraging their cultural and linguistic heritage for optimum survival, a process that yields linguistic and cultural hybridity.

The complexity that underlies the African immigrants’ linguistic and cultural identities in the United States should inform approaches to developing culturally competent education for youth and their families to improve overall health outcomes. It is expedient to target cancer-awareness information by incorporating aspects of the home languages of African immigrants—especially Western and Eastern Africa [5]. For example, the Swahili language would be accessible to immigrant families of East African origin, and Pidgin English for families with West African heritage. Appropriate learning mediums for cancer awareness for African immigrant youth should intersect with the linguistic and cultural practices of the African immigrant population.

**Belief System and Religious Affiliation**

In a 2021 report, the Pew Research Center stated that African immigrants in the United States are more religious than other Black Americans, even though Black Americans are more religious than Americans of other races [81]. Further broken down into specific practices, the report noted that African immigrants value attending religious services weekly, more than other Black Americans: “around half of the African immigrants living in the United States (54%) say they attend religious services at least weekly, compared with about 3-in-10 United States-born (32%) and Caribbean-born (30%) Black adults.”

Similar to language, culture, and national consciousness, the belief systems and religiosity of African immigrants will have a major imprint on their young children. Health information tailored specifically to religiosity will not only be responsive to African immigrants’ cultural perspectives, but it may also improve engagement with pedagogical materials. Moreover, studies are scarce on the intersection of African immigrants’ religious practices and responses to health care education about cancer, thereby illustrating another gap in research that may ultimately improve the approach to cancer education among distinctive ethnic and racial groups. Additionally, there is a shortage of research on the religious practices of African immigrants, highlighting another research gap that could ultimately enhance approaches to cancer education among distinct ethnic and racial groups.

**Hybrid Nationality and Afropolitanism**

African immigrants in the United States, have diverse origins from one of the 54 nations of Africa, many of which are multiethnic. These diverse ethnic identities house unique cultural and linguistic features within and outside the individual nation’s borders. While African immigrants actively engage with the dominant Western traditions of the society they reside in, they also maintain their cultural customs. As a result, youth from African immigrant families often exhibit hybrid language use, blending the host language with elements of African culture, including specific exclamation and colloquial forms rooted in African cultural beliefs. This linguistic and cultural hybridity is significant in addressing the existing gap in cancer awareness research among African immigrant families and fosters a sense of community within the African immigrant population in the United States.

The concept of Afropolitanism defines Africans as an integral part of the global community rather than separate from it. This concept refers to the empowerment associated with a blended, polyethnic, and cosmopolitan identity [80]. Afropolitanism iterates Africans’ awareness of their origins and the consciousness of the cultural ambiguities that occur because of their integration into the host society. This understanding impacts African immigrants’ response to cancer education and approaches to accessing health care for cancer treatment. Their cultural and spiritual beliefs are still very much prominent in their perspective on cancer disease. This consciousness could, however, be tapped into for possible changes and adaptations among this immigrant group. The summary of the African immigrants’ complex experience is iterated in the term, “Afropolitan.” Afropolitan describes an individual whose identities are deeply rooted in their diverse, transcultural experiences, reflecting youth linguistic and cultural practices within African immigrant families [68,69]. African immigrants’ hybrid language and cultural identities necessitate the development of health educational tools and technologies that integrate African cultural perspectives and engage these youth in learning and retaining health information in a culturally responsive manner.

**Age-Related Learning Preferences**

Consideration for age-appropriateness in technology is not unique to African immigrant youth; however, the connection of this factor to digital literacy, access, and equity makes it critical to examine further and worthy of discussion. A report by the Migration Policy Institute on immigrant learning with digital technology has identified uneven access to digital resources for youth aged between 15 and 17 years who are either
immigrants themselves or have at least one immigrant parent [82]. Research suggests that factors like work, language, and familial influence affect how immigrant youth use technology for learning [83]. Given the versatility of the adolescent age group with technology, they have increased access to vital information on health issues and diseases that are prevalent within their community. More important is their access to their heritage culture as well as the culture of their residing society. As a result, youth play a vital role as intermediaries, connecting with their families to promote cancer awareness within their communities.

Textbox 1. Summary of key main areas for future research.

<table>
<thead>
<tr>
<th>Priorities for future cancer prevention and control research focused on African immigrant populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Disaggregate study populations according to country or region of origin to improve cultural tailoring and knowledge.</td>
</tr>
<tr>
<td>• Develop family-oriented educational initiatives including programs for children.</td>
</tr>
<tr>
<td>• Use community-engaged approaches including partnerships with faith-based organizations.</td>
</tr>
<tr>
<td>• Leverage emerging technology for recruiting study participants and delivering educational messages while accounting for barriers to access.</td>
</tr>
<tr>
<td>• Align cancer awareness information with language and cultural markers specific to the population.</td>
</tr>
<tr>
<td>• Consider the global African community and hybrid African and American cultural practices.</td>
</tr>
<tr>
<td>• Incorporate relevant religious and spiritual beliefs and practices to enhance cancer education effectiveness.</td>
</tr>
<tr>
<td>• Consider youth and adolescents as intermediaries for increasing cancer awareness among family members.</td>
</tr>
<tr>
<td>• Explore the potential for interagency collaboration (Centers for Disease Control and Prevention, Centers for Medicare &amp; Medicaid Services, Health Resources and Services Administration, and National Institutes of Health) to address cancer-related health challenges for African immigrant families.</td>
</tr>
</tbody>
</table>

To achieve the goal of increasing cancer awareness among African immigrant families, 1 strategy involves creating a culturally tailored serious game. Serious games offer opportunities to build upon the research base of effective approaches to reduce the cancer burden by focusing on youth and leveraging technology. Research is crucial that examines the language use of youth from African immigrants in the United States. Previous research has already categorized most African immigrants living in the United States into Western (35.71%) and Eastern Africa (29.61%) groups, which could serve as a basis for examining youth cancer awareness within each group [5]. Open-ended ethnographic interviews could be used to identify the nuanced cultural and linguistic practices of the youth of African immigrant families. The heterogeneity of Africa’s cultural identities could result in a new monolithic idea of Black subgroups in the United States, the importance of beginning this inquiry cannot be delayed. Detailed demographic questionnaires and open response forms can allow for flexibility in how studies aggregate and allow for new divisions and aggregations of African immigrants. However, it is noted that migration by African countries is unequal with many African immigrants arriving from Western and Eastern African countries [5]. Additionally, recruitment strategies are particularly important in the success of this line of research and will need to be evaluated. As immigrant populations are “Hard-to-Reach,” using innovative ways to reach a target population is also important [84]. During the COVID-19 pandemic, online recruitment using Facebook (Meta), Instagram (Meta Platforms), and WhatsApp (WhatsApp LLC) was an effective recruitment strategy especially because it built on existing communication and information-sharing norms within the African immigrant community. Further research should use and evaluate multiple recruitment streams.

Findings from such research endeavors will have a meaningful impact on the strategies for developing culturally tailored educational content such as a serious game, to create awareness about cancer among African immigrant families in the United States. A culturally adapted serious game has immense potential to be instrumental in improving awareness and cancer prevention strategies in African immigrant families.

Conclusion

The importance of culturally tailored cancer education for African immigrants is underscored by uncertainty. Issues surrounding the measurement of cancer prevalence in African immigrant populations exacerbate the uncertainty of how cancer affects the African immigrant population in the United States. The existing, yet limited research on the topic suggests that African immigrants, especially adolescents, have unique experiences that lie at the nexus of their traditional culture and the complex novelty of the US health care system for immigrants. Research demonstrates the impact of cultural beliefs (such as fatalism and stigmatization of cancer diagnoses among African cultures) and lack of knowledge about cancer and cancer screening compounds to affect access to screening and care for African immigrants. Further research specifically targeting African immigrants and their youth can not only disentangle...
the unique struggles of African immigrants but also allow for the tailoring of education to provide maximal impact to vulnerable populations.

While recognizing our lack of knowledge and the uncertainty around the experience of cancer in the United States for African immigrants and advocating that increased research is the needed foundation for alleviating health disparities, more difficult work is ahead. It is integral for health scientists, health care providers, African culture scholars, and communities of African immigrants to come together for sustained research activity. These transdisciplinary associations will aid in the collection of data specific to African immigrants, but also the community engagement needed to co-design a culturally sensitive educational intervention. This will be no small task and require the dedicated work of many experts alongside and within the African immigrant community to forge long-term relationships that can facilitate recruitment, retention, and meaningful knowledge generation for the African immigrant community in the context of cancer experience.

Acknowledgments

Generative artificial intelligence technology was not used in the development or editing of this paper.

Authors' Contributions

OA conceptualized the original idea for this paper, led the gathering of data and team expertise, and contributed to drafting and leading the revising of the final paper. All coauthors contributed to the literature review and drafting of this paper.

Conflicts of Interest

None declared.

References


Abbreviations

**CRP:** culturally relevant pedagogy
An Introduction to the OutSMART Cancer Serious Game: Current and Future Directions

Olufunmilola Abraham¹, BPharm, MS, PhD; Tyler J McCarthy¹, BA
Social and Administrative Sciences Division, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Olufunmilola Abraham, BPharm, MS, PhD
Social and Administrative Sciences Division
University of Wisconsin-Madison
Rm 2515 Rennebohm Hall
777 Highland Ave
Madison, WI, 53705
United States
Phone: 1 6082634498
Fax: 1 6082625262
Email: olufunmilola.abraham@wisc.edu

Abstract

Given that cancer is a challenging disease that plagues millions of individuals of all age groups and socioeconomic statuses globally, developmentally appropriate education is often lacking for young people, particularly adolescents. Increasing cancer awareness and prevention education among adolescents using innovative strategies, such as game-based learning, is critical in reducing the burden of this disease. Adolescents are understudied in the field of cancer prevention and control, yet vulnerable as they tackle creating life-long health behavior patterns. Targeting cancer prevention education for adolescents has the potential to support long-term healthy behavior and reduce their risk of cancer. This paper provides an overview of the Collaborative Research on MEdication use and family health (CRoME) Lab’s novel game-based cancer prevention education tool. OutSMART Cancer is an innovative, novel educational intervention in the form of a serious game. Serious games are educational tools that seek to impart knowledge and improve behaviors in their players. This game covers information related to breast cancer, colon cancer, and lung cancer. This viewpoint is a summary of the developmental process for the OutSMART Cancer game. We describe in detail the work preceding initial game development, the current version of the game, future directions for the game, and its educational potential. The long-term goal of OutSMART Cancer is to improve cancer awareness and knowledge regarding prevention behaviors in adolescents and support a lifetime of health and wellness.

Introduction

The ubiquity of cancer in the United States denotes both the immense burden of the disease and the countless individuals devoted to spreading awareness. Although preventative and treatment-based measures have improved outcomes and reduced cancer deaths, the incidence of some of the most common cancers is on the rise [1]. Some studies have suggested that people in the United States lack the necessary knowledge and awareness of cancer, particularly those of lower socioeconomic status [2,3]. Similar results have been indicated in adolescents in the United Kingdom and the United States [4-7]. Thus, despite the prevalence of cancer awareness, there is still a salient and critical need to encourage cancer awareness and knowledge from a young age so that individuals can better understand the basic biological etiology of this disease and support life-long prevention behaviors.

For educational interventions to be successful, especially those involving complex and emotionally charged chronic conditions, they must be tailored to the intended audience. The Collaborative Research on MEdication use and family health (CRoME) Lab has a history of engaging with and educating...
adolescents and parents on health topics, such as cancer prevention, medication safety, and vaping prevention [8]. The CRoME Lab has co-designed serious games for adolescents and parents around prescription opioid medication safety [9]. Serious games are games designed with the characteristic purpose of imparting knowledge to the player, rather than merely providing a recreational experience [10]. In this viewpoint, we detail the work leading to the development of a cancer education serious game named OutSMART Cancer, the current state of the game, and future directions.

**Early Work With Youth Stakeholders**

In 2020, our team began research in Wisconsin with middle and high schools, holding focus groups with 327 students and conducting a further survey with 235 students [11-13]. In one study, the CRoME Lab held 25 focus groups with 188 middle and high school students between the ages of 12 and 18 years [11]. Through exploring adolescent perceptions, we found that many adolescents were interested in learning about cancer, specifically, cancer prevention. Middle and high school students in this study recounted familiarity with basic cancer biology but indicated unfamiliarity with how to assess cancer risk and what behaviors they can institute to prevent cancer.

One survey study examined adolescent’s knowledge and attitudes toward cancer as well as the acceptability of a game-based learning approach for cancer education in homes, health care settings, and schools [12]. The survey responses reiterated the findings from the initial previous focus groups. Although most students expressed basic cancer knowledge, only 66% knew that individuals have some level of control over their cancer risk. Moreover, only 37.3% reported knowing how to lower their cancer risk, while 50% suggested they try to make healthy choices to reduce their risk. Study findings provided further evidence for the need to educate youth on cancer and its prevention. Most adolescents (82%) reported that they would accept the use of a game to help them learn about cancer.

These initial studies with adolescents informed the CRoME Lab’s design of the OutSMART Cancer gameplay book, which was further assessed through focus groups with adolescents. This gameplay book showed adolescents the initial conceptualization of OutSMART Cancer informed by the Cancer Clear and Simple Curriculum [14]. A total of 18 focus groups, comprising 139 adolescents, provided in-depth feedback on the playbook [13]. Adolescents indicated that they preferred a serious game over educational modalities, such as websites and videos. Our cumulative research to date has shown that a serious game over educational modalities, such as websites and books, is more effective [15]. Adolescents indicated that they preferred a serious game over educational modalities, such as websites and videos. Our cumulative research to date has shown that a serious game over educational modalities, such as websites and videos, is more effective in improving adolescent cancer prevention behavior.

In 2023, OutSMART began early evaluation by adolescents and parents. This demonstration is currently unavailable to the public, as it is evaluated among key stakeholders—adolescents and parents. Informed consent has been collected from participants in each study related to the development and testing of this game. Findings from this study will result in an adapted version of the game, which will be used to evaluate efficacy and implementation.

**The OutSMART Cancer Game**

OutSMART is a web-based, computer videogame that presents 3 familiar, cancer-related scenarios in a narrative, choice-based format (Figures 1-5) [15]. It is built upon the Unity WebGL game engine and is currently optimized for browser gameplay on laptops and computers [16]. Within this game, players interact with the environment through a first-person perspective, taking the role of an adolescent. Players are faced with 3 scenarios that cover information related to breast, colon, and lung cancers. In each scene, players progress by clicking on pop-up bubbles, giving the player choices that move the storyline along. Players progress from one scene to the next once they have completed that scene’s storyline. After completion of a scene, players are taken to a map to choose the next available scene.

In the first scene, the player heads downstairs for a day at school to see their mother on the couch. After asking why she had not left for work, the mother tells the player that she is experiencing a painful lump in her breast alongside fatigue and will be going to her doctor that day. As the player offers to attend the appointment with her, the scene switches to the car ride to the doctor’s office. During this ride, the mother shares her family history of breast cancer and her anxiety. Later, in the doctor’s office, the player and their mother learn that the lump is cancerous. This level espouses key information, such as early warning signs of breast cancer, screening tools, basic cancer biology, and cancer stages.

In the second scene, the player brings mail inside for their father and discovers a letter from his doctor’s clinic encouraging him to schedule a colonoscopy. During this scene, the player tells their father about the letter and encourages him to schedule an appointment. However, at first, their father is apprehensive and talks of anxiety after his previous colonoscopy, which had uncovered polyps. After some conversation with their father, his attitude changes, and he decides to schedule a visit. This scene introduces colon cancer and screening strategies; it also introduces the player to the types of emotions that can act as barriers to screening.

In the final scene, the player goes to school and learns about cancer in a simulated classroom environment. Following class, their friend introduces them to a new person, who, when left alone with the player offers them a vape (electronic cigarette). Although peer pressure is evident, the player must refuse, articulating their own reasoning why they are choosing to protect their own health. This scene introduces players to the power of personal choice and how everyday choices can influence cancer risk.
Figure 1. OutSMART Cancer gameplay.

Figure 2. Screenshot of mom lying ill on the couch.

Figure 3. Screenshot of the doctor’s office scene.

Figure 4. Colon cancer scene with father.
The Future of the OutSMART Cancer Game

Although preliminary evaluation and targeted adaptation of this game are ongoing, a simultaneous endeavor has begun to create a more adapted cancer-education game. This adaptation of OutSMART Cancer will emphasize the need for cancer awareness and targeted education among Black Americans and African immigrant youth and parents (OutSMART Cancer: Africana). The creation of OutSMART Cancer: Africana is responsive to the need for culturally competent cancer education for youth and African immigrant families living in the United States using culturally familiar language and imagery. Black Americans and African immigrants experience cancer and health care uniquely, requiring a tailored educational approach [17].

The initial intention is to widely disseminate this serious game through clinical settings, such as community pharmacies, primary care offices, and cancer clinics, as well as community settings, such as schools and community health organizations. Researchers aim for this intervention to be taken up by adolescents and their families across the United States. We expect that this game will eventually be publicly available on the internet for play on computers and mobile devices. The long-term goal of the OutSMART Cancer games is to facilitate family communication about cancer prevention, associated healthy behaviors, early detection, and cancer screening.

Authors' Contributions

OA conceptualized and supervised the study and was in charge of writing the original draft as well as reviewing and editing the manuscript. TJM contributed to the writing of the original draft as well as reviewing and editing it.

Conflicts of Interest

None declared.

References


15. OutSMART Cancer. URL: https://crome.wisc.edu/OutSmart/#top [accessed 2023-11-13]


Abbreviations

CRoME: Collaborative Research on MEdication use and family health

©Olufunmilola Abraham, Tyler J McCarthy. Originally published in JMIR Cancer (https://cancer.jmir.org), 29.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Iterative Patient Testing of a Stimuli-Responsive Swallowing Activity Sensor to Promote Extended User Engagement During the First Year After Radiation: Multiphase Remote and In-Person Observational Cohort Study

Eileen H Shinn¹, PhD; Adam S Garden², MD; Susan K Peterson¹, PhD; Dylan J Leupi³, BS; Minxing Chen³, MS; Rachel Blau⁴, PhD; Laura Becerra⁶, MS; Tarek Rafeedì⁵; Julian Ramirez⁵, PhD; Daniel Rodriquez⁵, PhD; Finley VanFossen⁴, MA; Sydney Zehner⁴, BS; Patrick P Mercier⁶, PhD; Joseph Wang⁵, PhD; Kate Hutcheson⁷, PhD; Ehab Hanna⁷, MD; Darren J Lipomi⁵, PhD

¹Department of Behavioral Science, University of Texas, MD Anderson Cancer Center, Houston, TX, United States
²Department of Radiation Oncology, University of Texas, MD Anderson Cancer Center, Houston, TX, United States
³Department of Chemistry and Biochemistry, College of Science, University of Notre Dame, South Bend, IN, United States
⁴Department of Biostatistics, University of Texas, MD Anderson Cancer Center, Houston, TX, United States
⁵Department of Nano and Chemical Engineering, University of California, San Diego, CA, United States
⁶Department of Electrical and Computer Engineering, University of California, San Diego, CA, United States
⁷Department of Head and Neck Surgery, University of Texas, MD Anderson Cancer Center, Houston, TX, United States

Corresponding Author:
Eileen H Shinn, PhD
Department of Behavioral Science
University of Texas
MD Anderson Cancer Center
1155 Herman Pressler
Unit 1330, PO Box 301439
Houston, TX, 77230-1330
United States
Phone: 1 713 745 0870
Fax: 1 713 745 4286
Email: eshinn@mdanderson.org

Abstract

Background: Frequent sensor-assisted monitoring of changes in swallowing function may help improve detection of radiation-associated dysphagia before it becomes permanent. While our group has prototyped an epidermal strain/surface electromyography sensor that can detect minute changes in swallowing muscle movement, it is unknown whether patients with head and neck cancer would be willing to wear such a device at home after radiation for several months.

Objective: We iteratively assessed patients’ design preferences and perceived barriers to long-term use of the prototype sensor.

Methods: In study 1 (questionnaire only), survivors of pharyngeal cancer who were 3-5 years post treatment and part of a larger prospective study were asked their design preferences for a hypothetical throat sensor and rated their willingness to use the sensor at home during the first year after radiation. In studies 2 and 3 (iterative user testing), patients with and survivors of head and neck cancer attending visits at MD Anderson’s Head and Neck Cancer Center were recruited for two rounds of on-throat testing with prototype sensors while completing a series of swallowing tasks. Afterward, participants were asked about their willingness to use the sensor during the first year post radiation. In study 2, patients also rated the sensor’s ease of use and comfort, whereas in study 3, preferences were elicited regarding haptic feedback.

Results: The majority of respondents in study 1 (116/138, 84%) were willing to wear the sensor 9 months after radiation, and participant willingness rates were similar in studies 2 (10/14, 71.4%) and 3 (12/14, 85.7%). The most prevalent reasons for participants’ unwillingness to wear the sensor were 9 months being excessive, unwanted increase in responsibility, and feeling self-conscious. Across all three studies, the sensor’s ability to detect developing dysphagia increased willingness the most compared
to its appearance and ability to increase adherence to preventive speech pathology exercises. Direct haptic signaling was also rated highly, especially to indicate correct sensor placement and swallowing exercise performance.

Conclusions: Patients and survivors were receptive to the idea of wearing a personalized risk sensor for an extended period during the first year after radiation, although this may have been limited to well-educated non-Hispanic participants. A significant minority of patients expressed concern with various aspects of the sensor’s burden and its appearance.

Trial Registration: ClinicalTrials.gov NCT03010150; https://clinicaltrials.gov/study/NCT03010150

(JMIR Cancer 2024;10:e47359) doi:10.2196/47359

KEYWORDS
user-centered design; patients with head and neck cancer; dysphagia throat sensor

Introduction

Background
In 2021, approximately 32,000 Americans developed laryngeal or pharyngeal cancer, which has a 5-year survival rate of 61% for all stages combined [1]. Management of these cancers often includes high-dose intensity-modulated radiation therapy (IMRT) designed to spare pharyngeal muscles and reduce the incidence of radiation-associated dysphagia (swallowing difficulty) [2]. Still, a range of studies have reported that roughly 60% of patients receiving IMRT developed long-term swallowing problems within 2 years after radiation had ended, ranging in intensity from inability to swallow solid food without compensatory strategies to being completely feeding tube dependent [3-10].

As with most chronic conditions, early detection and intensive swallowing therapies are key to preventing long-term dysphagia [11-26], especially if patients are adherent to swallowing therapy instructions [27]. However, noninvasive screening procedures for early detection of radiation-associated fibrosis do not yet exist in the United States. Instead, gold standard modified barium swallow (MBS) and fiberoptic endoscopic evaluation of swallowing tests are typically ordered after the patient begins to complain of difficulties with swallowing [12]. Furthermore, preventive swallowing therapies are not always prescribed prior to the development of radiation-associated dysphagia [28-30]. Unfortunately, once radiation-associated dysphagia is clinically detected, there is little hope of fully restoring normal function [11,31,32].

To detect radiation-associated dysphagia before it becomes permanent, it is necessary to monitor changes in swallowing function much more frequently than is currently possible in the clinical setting. Subclinical change in swallowing activity or risk for dysphagia could be assessed during standard cancer surveillance visits, but increasing the periodicity of these visits would increase patient burden by requiring more frequent travel to the medical center for swallowing imaging and tests. Frequent at-home monitoring with wearable sensors between scheduled surveillance visits could address this gap in monitoring, especially if the sensors were designed to support decision-making regarding initiation of intensive speech language therapies [33]. To this end, researchers have developed myriad devices that can be worn on the skin and measure a range of mechanical, optical, biochemical, electrical, or acoustic signals with high fidelity [34-39].

However, sensor performance alone is not sufficient for improving health outcomes as patient engagement is also important [40]. Within the specific context of preventing dysphagia in survivors of head and neck cancer, repeated at-home monitoring over a period of months if not years is necessary to demonstrate a clinical advantage over current treatment paradigms. Unfortunately, most mobile technologies fail to engage patients over sustained durations, with most mobile health (mHealth) interventions for chronic disease reporting steep declines in user engagement, some as high as 95% within the first few weeks, depending on the technology and context [41-43]. The most frequently cited reasons for discontinued use are decreased interest in the technology after its novelty abates, perceived lack of usefulness relative to burden, poor implementation of user experience, and frustration with technical issues [44-47].

To counter these barriers, it is widely agreed that user-centered testing be conducted in a sustained and iterative fashion during the design and development of new health technology. User-centered testing assesses the human technology interface by evaluating how well the technology incorporates into end users’ routines, habits, and capabilities, known loosely as user acceptance [40,48]. Beyond acceptability, technologies should be designed to maximize their potential to effect changes in patients’ attitudes and health behaviors. Oinas-Kukkonen and Harjumaa’s [49] persuasive system design model describes four categories of persuasive design principles that optimize the likelihood of health behavior change: task support (personalized design features that make it easier for users to achieve their goals), social support (leveraging interpersonal learning, eg, via online community forums), dialogue support (providing feedback to the user in a manner that helps the user move toward their goal, eg, with praise and rewards), and system credibility (the perceived clinical expertise embedded within the sensor output) [49]. Relatively few mHealth interventions conduct user-centered testing during technology development, which may be one reason for diminishing patient engagement and eventual abandonment [50-53]. In the US market, the user abandonment rate of fitness trackers is 50% within 6-12 months [44,54]. Patient abandonment rates are higher for those 70 years and older; one study found that 43% of their sample had abandoned their sleep and activity trackers within the first 2 weeks of use [55].

A recent review of 51 mHealth intervention studies targeting chronic diabetes, cardiovascular, or pulmonary diseases noted that diminished patient engagement was prevalent and posed a
significant threat to effective use of the technology. Accordingly, nonsignificant effects on clinical markers outweighed significant findings two-to-one [42]. Therefore, our study explicitly addressed the design of a wearable sensor with the future intended use of home-based assessment for 9 months, starting with the third month after radiation to the 12th month. All design preferences and opinions were solicited within the context of sustaining engagement with the sensor for 9 months during the first year since repeated measurements over time would be needed to detect patterns of developing dysphagia in posttreatment patients.

Goal of This Study
We assessed patient needs and preferred characteristics regarding the design of a wearable sensor to deliver personalized risk of dysphagia. Specifically, we assessed perceived barriers to wearing the sensor for 9 months, starting from the third month after the end of radiation treatment (to allow for healing from radiation skin burn) until the 12th month post treatment, and the impact of proposed design features on willingness to wear the sensor. In the first of three iterative user-centered tests, we surveyed a large cohort of survivors of head and neck cancer who were 4-5 years past radiation treatment to assess the perceived need for the sensor and desired design features for future prototypes. In study 2, we assessed user acceptability for a wired prototype sensor within a small sample of long-term survivors, oversampled for radiation-associated dysphagia. Finally, in the third user test, we tested a revised prototype on a sample of patients with head and neck cancer undergoing active treatment to get a better sense of competing priorities during a fraught time in their lives. The revised prototype included more elastic and comfortable materials for the strain sensor and custom-made dry electromyography (EMG) sensors, as opposed to commercial sensors. During the third test, we repeated our questions about user acceptability and willingness to wear the sensor for 9 months, as well as new questions about bidirectional feedback in the form of haptic (vibration) signaling.

Methods

Study 1

Design and Eligibility
Survivors of head and neck cancer who were still alive and who were already enrolled in a psychosocial parent study were asked to answer a questionnaire about a hypothetical throat sensor. Men and women were eligible for the parent study if they had received radiation with curative intent for oropharyngeal (stage II-IVb), laryngeal (II-IVb), hypopharyngeal (I-IVb), or nasopharyngeal cancer (I-IVb), or an unknown primary cancer with cervical metastases; were at least 2 years post treatment; were 18 years or older; and spoke English. Men and women were excluded if they had treatment for previous head and neck cancer; a history of previous head and neck surgery (previous biopsy, tonsillectomy, or tracheotomy were allowed); other cancer diagnoses, except nonmelanoma skin cancer; or a history of current oropharyngeal dysphagia unrelated to cancer diagnosis (eg, dysphagia due to underlying neurogenic disorder).

Recruitment and Data Collection Procedures
For the psychosocial parent study, all eligible patients were approached for recruitment at the radiation clinic’s radiation education class after being identified at the weekly multidisciplinary tumor board conference. The accrual rate for entry into the original parent study was 77%; demographic and disease information was collected at baseline. Those patients who were already enrolled in the psychosocial parent study and still alive (n=234) were contacted by phone to determine if they would answer optional questions about a hypothetical sensor to be worn on the throat. Patients who did not return calls after 5 attempts or did not have working phone numbers were not approached further for enrollment into study 1. After obtaining informed consent, participants completed the optional questionnaire administered either by REDCap, telephone, or mail at a single time point [56]. For mailed questionnaires, a research staff’s phone number was provided if the patient had any questions about the questionnaire.

Measures
Demographic information regarding age, race/ethnicity, employment, income, and marital status were obtained by questionnaire. Disease stage was abstracted from the medical record. Participants then completed a questionnaire. The first page of the questionnaire showed a photograph of the proposed sensor (Figure 1A) and a diagram of the sensor’s placement on the neck (Figure 1B), a brief description of the sensor’s purpose, and the proposed timeline of wearing the sensor every weekend from the third month post radiation to the 12th month post radiation for a total of 9 months.
Figure 1. Appearance of the hypothetical and actual sensor prototypes. (A) Study 1 respondents were shown a photograph of the proposed sensor and (B) its proposed location on the neck. (C) Study 2’s graphene strain sensor prototype, supported on 13-µm-thick polyimide tape (the contact surface is silicone) and placed on the submental region probing muscle contraction. (D) Study 3’s soft polymer strain sensor, now placed under the laryngeal prominence to capture movement during swallowing.

Main Outcome: Willingness to Wear the Sensor

For studies 1-3, the study questionnaire asked whether the patient would have been willing to wear the sensor for 9 months during the first year after radiation, starting in month 3 post treatment. This time point was asked about since it would give sufficient time for the skin on their neck to have healed from radiation skin burn. Participants were then asked whether they would have been willing to wear the sensor for the entire 9 months, every other week, or every weekend during the 9-month period, and then a series of branched logic true-false questions about reasons for willingness versus unwillingness to wear the sensor. Next, using a 3-point Likert scale response format, all participants rated whether changes in the sensor design (either unobtrusive appearance or the ability to receive feedback about risk for dysphagia) would change the individual’s willingness or unwillingness to wear the sensor every weekend for 9 months. Additional comments or suggestions were also solicited as free text.

Study 2

Design and Eligibility

A second sample of survivors of head and neck cancer who were 2 to 10 years post radiation and attending surveillance visits at MD Anderson gave informed consent and enrolled into the study during a 1-week period; testing was constrained to a 1-week period in which visiting graduate engineering students from the University of California San Diego traveled to MD Anderson for on-patient equipment testing. The eligibility criteria for study 2 were the same as for study 1; however, we oversampled for patients with a Dynamic Imaging Grade of Swallowing Toxicity score >0, indicating radiation-associated dysphagia that had been verified with MBS [57]. The oversampling was done to gauge the accuracy of the prototype sensor in distinguishing between dysphagic survivors and survivors without dysphagia [58]. For every dysphagic participant, we recruited a nondysphagic patient matched for age and sex. For patients who declined participation, deidentified disease information, demographics, and reason for refusal were noted in the study record.

Procedure and Assessment

A wired prototype graphene strain sensor coupled with a wired surface EMG sensor was placed on the patient to obtain muscle movement measurements during a series of swallowing tasks of various bolus textures, as described previously (Figure 1C) [58]. Immediately after the on-throat sensor test, patients were asked to answer six questions about the sensor’s discomfort, ease of use, and associated embarrassment using a 5-point Likert scale ranging from strongly disagree to strongly agree. Patients
were again asked whether they would be willing to wear the sensor for 9 months (but now for once a month on the weekends) with branching questions asking for reasons for willingness versus unwillingness. Patients were again asked to rate the impact of sensor unobtrusiveness and predictive dysphagic feedback on willingness to wear the sensor for extended periods. Finally, demographic information regarding age, race, and marital status were abstracted from the medical record. All testing sessions were conducted at the Head and Neck Cancer Center at MD Anderson.

**Study 3**

**Design and Eligibility**

Similar eligibility, consent, and testing procedures were used in study 3. However, eligible patients were more likely to be approached during active treatment for throat cancer, whereas studies 1 and 2 recruited long-term survivors. Study 3’s sensor (Figure 1D) was revised to have better skin conformation and comfort; standard surface EMG electrodes were now replaced with flexible custom dry electrodes, whereas the strain sensor was supported on a silicon substrate [39].

**Assessment Procedures**

After completion of the on-throat sensor test, patients were also asked the same questions used in study 1 regarding willingness to wear the sensor for 9 months and whether changes in the sensor’s appearance and feedback capability would change their minds about their willingness to wear the sensor. In addition, participants were interviewed regarding the helpfulness of future capability of the sensor itself to give immediate haptic feedback in three different scenarios: to indicate correct placement of the sensor, to indicate correct performance of a particular swallowing exercise, and to indicate quality of swallowing during at-home testing of various bolus textures. Their answers were transcribed, categorized, and coded into three categories (0: not helpful; 1: helpful under certain conditions; 2: helpful).

**Analysis**

Descriptive statistics (eg, proportions, means, ranges, and SDs) were computed for the process evaluation and participant satisfaction data, together with 95% CIs. To assess the external validity of the study, demographic and disease information was compared between respondents and nonrespondents in study 1 (Table 1) and between participants and refusers in studies 2 and 3 (data not shown; available data for participants in studies 2 and 3 is shown in Table 2). All questionnaire responses were analyzed with SPSS (version 26; IBM Corp).
**Table 1.** Demographic/disease comparisons between willing and unwilling participants (study 1).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Potentially eligible survivors (from parent study)</th>
<th>Survivors who completed the questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total (N=234) Nonrespondent (did not participate)</td>
<td>Respondent</td>
</tr>
<tr>
<td>What is your age? (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n (%)</td>
<td>234</td>
<td>96 (41.0)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>57.4 (10.0)</td>
<td>56.6 (9.8)</td>
</tr>
<tr>
<td>Median (min-max)</td>
<td>58 (18-83)</td>
<td>56 (30-79)</td>
</tr>
<tr>
<td>What is your ethnic background? n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>21 (9.1)</td>
<td>15 (16.0)</td>
</tr>
<tr>
<td>Non Hispanic or Latino</td>
<td>210 (90.9)</td>
<td>79 (84.0)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>10 (4.3)</td>
<td>6 (6.4)</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1 (0.4)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td>Asian</td>
<td>6 (2.6)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander</td>
<td>1 (0.4)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>213 (92.2)</td>
<td>86 (91.5)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college and lower</td>
<td>112 (48.9)</td>
<td>54 (58.1)</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>117 (51.1)</td>
<td>39 (41.9)</td>
</tr>
<tr>
<td>Employment status, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time/part-time</td>
<td>145 (63.3)</td>
<td>57 (61.3)</td>
</tr>
<tr>
<td>Not employed</td>
<td>84 (36.7)</td>
<td>36 (38.7)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single living alone/married but living apart/separated/divorced/widow</td>
<td>46 (20.0)</td>
<td>21 (22.1)</td>
</tr>
<tr>
<td>Single but living with significant other/married living with spouse</td>
<td>184 (80.0)</td>
<td>74 (77.9)</td>
</tr>
<tr>
<td>Occupation, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional/managerial</td>
<td>143 (71.9)</td>
<td>51 (63.0)</td>
</tr>
<tr>
<td>Retail/service/labor</td>
<td>44 (22.1)</td>
<td>24 (29.6)</td>
</tr>
<tr>
<td>Student/unemployed</td>
<td>12 (6.0)</td>
<td>6 (7.4)</td>
</tr>
<tr>
<td>What is your income before taxes? (US $), n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30,000</td>
<td>38 (18.9)</td>
<td>24 (30.4)</td>
</tr>
<tr>
<td>30,000-50,000</td>
<td>31 (15.4)</td>
<td>13 (16.5)</td>
</tr>
<tr>
<td>50,000-75,000</td>
<td>28 (13.9)</td>
<td>9 (11.4)</td>
</tr>
<tr>
<td>&gt;75,000</td>
<td>104 (51.7)</td>
<td>33 (41.8)</td>
</tr>
<tr>
<td>Stage of disease, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stages I or II</td>
<td>76 (32.5)</td>
<td>27 (28.1)</td>
</tr>
<tr>
<td>Stages III or IV</td>
<td>158 (67.5)</td>
<td>69 (71.9)</td>
</tr>
<tr>
<td>Characteristic</td>
<td>Study 2</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td></td>
<td>Total sample</td>
<td>Willing to wear for 9 mo</td>
</tr>
<tr>
<td></td>
<td>(n=14)</td>
<td>(n=10)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>61.6 (11.5)</td>
<td>61.2 (12.3)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>American Indian or Alaskan Native</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Asian</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>13 (92.9)</td>
<td>9 (90)</td>
</tr>
<tr>
<td>More than one race</td>
<td>1 (7.1)</td>
<td>1 (10)</td>
</tr>
<tr>
<td>What is your ethnic background? n (%)</td>
<td>.85</td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>3 (21.4)</td>
<td>2 (20)</td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>11 (78.6)</td>
<td>8 (80)</td>
</tr>
<tr>
<td>Occupation, n (%)</td>
<td>.52</td>
<td></td>
</tr>
<tr>
<td>Managerial/professional</td>
<td>2 (14)</td>
<td>2 (20)</td>
</tr>
<tr>
<td>Retail, service, operator</td>
<td>9 (64)</td>
<td>6 (60)</td>
</tr>
<tr>
<td>Student or unemployed</td>
<td>3 (21)</td>
<td>2 (20)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td>.37</td>
<td></td>
</tr>
<tr>
<td>Married/living with significant other</td>
<td>12 (86)</td>
<td>8 (80)</td>
</tr>
<tr>
<td>Single/divorced/widowed/separated</td>
<td>2 (14)</td>
<td>2 (20)</td>
</tr>
<tr>
<td>Dysphagic status, n (%)</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Dysphagic (DIGESTa&gt;0)</td>
<td>7 (50)</td>
<td>5 (50)</td>
</tr>
<tr>
<td>Not dysphagic (DIGEST=0)</td>
<td>7 (50)</td>
<td>5 (50)</td>
</tr>
<tr>
<td>Disease stage, n (%)</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>I-II</td>
<td>1 (8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>III-IV</td>
<td>12 (92)</td>
<td>9 (90)</td>
</tr>
</tbody>
</table>

*aDIGEST: Dynamic Imaging Grade of Swallowing Toxicity.

bN/A: not applicable.

Ethical Considerations

All study materials and procedures were approved by the institutional review board at MD Anderson Cancer Center's institutional review board (protocol 2016-0597). All enrolled participants signed informed consent forms before testing began. All study data were deidentified, and no compensation was provided for participation.

Results

Overview

Prior to patient user testing, our study incorporated design input from multiple disciplines, including behavioral scientists, speech pathologists, radiation oncologists, and engineers. Initially, our primary concerns were to develop a wearable device that would not injure skin sensitized by radiation and have an uncomplicated application and removal procedure. Various invasive sensors, such as those worn inside the mouth, were dropped from consideration after it was realized that patients would possibly need to use the device during radiation and later at home during the first year post treatment. During study 1, we gathered patient reactions to a photograph of a sensor (Figure 1), whereas in studies 2 and 3, prototype versions were tested on survivors and patients in the clinic (Figure 1). The racial breakdown of the overall study sample (N=234) was non-Hispanic White (n=213, 92.2%), African American (n=10, 4.3%), Asian American (n=6, 2.6%), American Indian/Alaska
Native (n=1, 0.4%), and Native Hawaiian/Pacific Islander (n=1, 0.4%).

Study 1
Research staff contacted 234 eligible participants to complete study 1’s questionnaire, either via REDCap or by mail; 138 (59%) participants completed the questionnaire (Figure 2). Participants in study 1 were primarily non-Hispanic White and married, and their mean age was 57.4 (SD 10) years (Table 1). Median time since end of radiation treatment was 4 years and 26 days (Table 1). Analyses of responders versus nonresponders showed that responders were more likely to be non-Hispanic, have a bachelor’s degree, and have higher annual income; differences in race, age, and disease stage were not significantly different (Table 1).

Figure 2. Recruitment CONSORT (Consolidated Standards of Reporting Trials) for study 1 (n=138).

Survivor Preferences for Wearable Throat Sensor
Of the 138 respondents, 115 (83.3%) agreed that they would have been willing to wear the sensor for 9 months during the first year after radiation. However, patients were not willing to wear the sensor during the workweek due to fear of coworkers or strangers asking about the sensor. Instead, they were willing to wear the sensor on weekends, but only for one weekend a month as opposed to every weekend. When presented with several potential reasons explaining their willingness to wear the sensor, nearly all participants cited altruism, whereas 88% (92/105) cited interest in the sensor technology and 77% (75/97) thought that the sensor would help them adhere to their preventive swallowing exercises (Table 3). For example, several patients commented that the personalized feedback from the sensor would provide additional motivation to adhere to their preventive swallowing exercises:

- It would push me to do my exercises diligently...
- It would get me on the ball and do my exercises more often...
- It would give me the information I can use to fight back the scar tissue problem. And see the importance of my neck exercises.
Table 3. Studies 1-3: Number of patients endorsing reasons for willingness/unwillingness to wear the sensor every weekend for 9 months.

<table>
<thead>
<tr>
<th>Reasons for willingness/unwillingness to wear the sensor for 9 months&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Would wear sensor, n (%)</th>
<th>Would not wear sensor, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>False</td>
</tr>
<tr>
<td><strong>Study 1 (n=138)&lt;sup&gt;b&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The technology of the patch sounds interesting.</td>
<td>92 (87.6)</td>
<td>13 (12.4)</td>
</tr>
<tr>
<td>Wearing the patch would have reminded me to do my swallowing exercises.</td>
<td>75 (77.3)</td>
<td>22 (22.7)</td>
</tr>
<tr>
<td>I wanted to help with MD Anderson’s research.</td>
<td>108 (99.1)</td>
<td>1 (0.9)</td>
</tr>
<tr>
<td>My skin was still sensitive during that time.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I wouldn’t want to put on and take off the patch every weekend.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I wouldn’t want to wear the patch for 9 months.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I would feel uncomfortable if people noticed the patch and ask me questions or wanted to talk about it.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I was being asked to participate in too many studies.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>It would have added to my daily responsibilities.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>It would have been a reminder of my cancer treatment.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I would not be able to see my data from the patch.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>Study 2 (n=14)&lt;sup&gt;d&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The technology of the patch sounds interesting.</td>
<td>8 (80)</td>
<td>2 (20)</td>
</tr>
<tr>
<td>Wearing the patch would have reminded me to do my swallowing exercises.</td>
<td>10 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>I wanted to help with MD Anderson’s research.</td>
<td>10 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>My skin was still sensitive during that time.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I wouldn’t want to put on and take off the patch every weekend.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I wouldn’t want to wear the patch for 9 months.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I would feel uncomfortable if people noticed the patch and ask me questions or wanted to talk about it.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I was being asked to participate in too many studies.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>It would have added to my daily responsibilities.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>It would have been a reminder of my cancer treatment.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I would not be able to see my data from the patch.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>Study 3 (n=14)&lt;sup&gt;f&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The technology of the patch sounds interesting.</td>
<td>12 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Wearing the patch would have reminded me to do my swallowing exercises.</td>
<td>12 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>I wanted to help with MD Anderson’s research.</td>
<td>12 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>My skin was still sensitive during that time.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I wouldn’t want to put on and take off the patch every weekend.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I wouldn’t want to wear the patch for 9 months.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I would feel uncomfortable if people noticed the patch and ask me questions or wanted to talk about it.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I was being asked to participate in too many studies.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>It would have added to my daily responsibilities.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>It would have been a reminder of my cancer treatment.</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>I would not be able to see my data from the patch.</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>a</sup>Participants were asked the following question “Which of the following reasons would motivate you to wear the sensor every weekend for 9 months after radiation?”
In study 1, 155 (83.5%) participants indicated that they would wear the sensor, while 23 (16.5%) participants indicated that they would not wear it.

Not applicable.

In study 2, 10 (71.4%) participants indicated that they would wear the sensor, while 4 (28.5%) participants indicated that they would not wear it.

In study 3, 12 (85.7%) participants indicated that they would wear the sensor, while 2 (14.3%) participants indicated that they would not wear it.

Others valued the additional information that the sensor would provide:

- I would be curious to know what is going on with my body...
- I would have liked to have known what was happening to my throat...
- It’s my neck! Why wouldn’t I want to know?

Among the 22 participants who indicated that they would have been unwilling to wear the sensor, nearly 90% (24/28, 85.7%) of all unwilling participants cited the lengthy duration of having to wear the sensor and 57% (16/28) disliked the idea of having to wear the sensor every weekend. The photograph of the proposed sensor had large black letters embedded within the sensor (Figure 1) to contain its wiring; over half of the unwilling participants objected to the sensor being noticeable enough that others would want to ask questions about its purpose. Just under one-third of unwilling participants disliked the idea of being reminded of their cancer treatment during the first year after radiation (Table 3). Participants who were unwilling to wear the sensor for 9 months did not have any significant demographic or clinical differences compared to participants who expressed willingness to wear the sensor. When asked whether changing the sensor’s appearance to that of a Band-Aid would impact willingness, 29% (4/14) of all study 1 participants agreed that this would increase their willingness, whereas 71% (10/14) stated that unobtrusive appearance would not affect their willingness (mean 2.45, SD 0.87; Figure 3):

Cosmetics is the least of my worries when I am going through treatment and fighting for my life.

When asked about the sensor’s proposed function of delivering individual risk for dysphagia, the majority of the sample (21/28, 75%) agreed that this feature would increase their willingness (mean 1.5, SD 0.88; Figure 3). Notably, half of the free-text comments indicated that had they been able to measure muscle fibrosis earlier, they would have been more diligent about performing their prescribed swallowing exercises. Some simply wrote that they wanted the sensor to be available so that future patients would understand that the risk of dysphagia was high:

I would like to see this in ACTION NOW

**Study 2**

Within a 1-week period, a convenience sample of 20 potentially eligible survivors of oropharyngeal cancer who were nonmetastatic and able to speak English were approached at their surveillance visit for enrollment into the study. To test the sensor’s performance in distinguishing between normal and dysphagic swallowing patterns, survivors who had developed severe dysphagia as a result of their radiation were oversampled for study 2. Potentially eligible survivors were first identified in the electronic medical record; approached during a surveillance visit; and if consented, scheduled with the engineers for the sensor testing session in a clinic exam room. Three patients refused to participate, citing fatigue or disinterest; all were White, 2 were male, and 1 was female, and their age ranged from 63 to 74 years. Two of the patients were dysphagic and the third was nondysphagic. All three had been diagnosed with late-stage oropharynx cancer (data not shown). A total of 17 (85%) patients agreed, but 1 patient subsequently dropped out due to receiving news of cancer recurrence (Figure 4). Another 2 participants experienced scheduling conflicts; informed consent was obtained from the remaining 14 participants.
Consistent with this cancer type’s demographic profile, the average age of the sample was 61 years, with 12 male participants and 2 female participants. Three participants were Hispanic or Latino and 3 were of non-White race (Table 2). Specific cancer diagnoses included cancer of the oropharynx (9/14, 64%), larynx (3/14, 21%), and nasopharynx (1/14, 7%), and unknown primary cancers (1/14, 7%). The average time since completion of radiation treatment was 47.9 months, and half of the sample had received a diagnosis of radiation-associated dysphagia (Table 2).

Figure 4. Recruitment flowchart for study 2 (n=14).

After wearing the sensor, 10 of the 14 (71%) patients indicated that they would have been willing to wear the sensor for 9 months of the first year post radiation. The most prevalent reasons for willingness were wanting to help future patients detect developing dysphagia and wanting to help MD Anderson research (Table 3). Of the 4 (29%) patients who did not think they would have been willing to wear the sensor, the most popular reason for unwillingness was study burden, specifically, that 9 months was too long of a testing period and the increased responsibilities associated with the sensor. Using a 5-point Likert response scale, patient ratings of discomfort (mean 1.21, SD 0.42), embarrassment (mean 1.14, SD 0.36), and difficulty in application and removal (mean 1.5, SD 0.52) were minimal (Table 4). Therefore, these questions were not repeated in the next phase of user testing.

Table 4. Study 2’s mean patient ratings for sensor discomfort, embarrassment, and difficulty of application (n=14), and study 3’s mean patient ratings of helpfulness for haptic signaling (n=14).

<table>
<thead>
<tr>
<th>Study 2 (n=14)</th>
<th>Patient ratings, mean (SD)</th>
<th>Range(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The sensor was uncomfortable to wear.</td>
<td>1.21 (0.426)</td>
<td>1.0-5.0</td>
</tr>
<tr>
<td>The sensor would be difficult for me to use at home.</td>
<td>1.5 (0.519)</td>
<td>1.0-5.0</td>
</tr>
<tr>
<td>I thought the experiment was fun.</td>
<td>3.79 (0.893)</td>
<td>1.0-5.0</td>
</tr>
<tr>
<td>The testing session was embarrassing.</td>
<td>1.14 (0.363)</td>
<td>1.0-5.0</td>
</tr>
<tr>
<td>I am good about doing my swallowing exercises every day.</td>
<td>3.27 (1.51)</td>
<td>1.0-5.0</td>
</tr>
<tr>
<td>I believe it is important for me to do as many of my swallowing exercises as possible.</td>
<td>4.46 (1.13)</td>
<td>1.0-5.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 3 (n=14)</th>
<th>Patient ratings, mean (SD)</th>
<th>Range(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would it help for the sensor itself to vibrate when you put it in the right spot on your throat?</td>
<td>1.85 (0.376)</td>
<td>0-2.0</td>
</tr>
<tr>
<td>Do you think it would be helpful to have the sensor vibrate once you did your swallowing exercise correctly?</td>
<td>2.00 (0.000)</td>
<td>0-2.0</td>
</tr>
<tr>
<td>Do you think that having the sensor process your swallowing data and give you feedback about the quality of your swallowing would help?</td>
<td>1.46 (0.877)</td>
<td>0-2.0</td>
</tr>
</tbody>
</table>

\(^a\)For study 2, the scale ranged from 1 (strongly disagree) to 5 (strongly agree). For study 3, the scale ranged from 0 (no) to 2 (yes).
Study 3
As with study 2, a convenience sample of 14 participants were recruited within a 1-week period to assess user preferences to the updated sensor prototype. As in the previous two studies, the majority of patients were diagnosed with oropharyngeal cancer (11/14, 79%). Unlike the previous two studies, 11 of the 14 (78.6%) were undergoing radiation at the time of testing; the remaining 2 participants were 1-5 year survivors (data not shown). The long-term dysphagic status was not yet known for patients on active treatment. A total of 17 participants were eligible and approached to participate in the sensor study. Two patients refused, both being White and male: 1 patient was aged 76 years and had been diagnosed with late-stage oropharyngeal cancer 2 years prior and the other was aged 23 years and was in the third week of radiation for late-stage oropharynx cancer (data not shown). A total of 15 (83%) patients agreed to participate and gave informed consent. One participant developed an acute illness episode the following day and was, therefore, unable to complete the sensor test, leaving 14 participants who completed user testing (Figure 5). Study 3’s sample was primarily male (12/14, 86%) and non-Hispanic White (12/14, 86%) with an average age of 62 years (Table 2).

Figure 5. Recruitment flowchart for study 3 (n=14).

As with the previous studies, the majority of patients (12/14, 86%) indicated willingness to wear the sensor for 9 months during the first year after radiation. Wanting to help future patients detect developing dysphagia and wanting to help MD Anderson research were the most prevalent reasons for willingness to wear the sensor (Table 3). As in study 2, the most oft-cited reasons for unwillingness were that of study burden (lengthy testing period and increase in daily responsibilities; Table 3). Patients’ opinions regarding the helpfulness of haptic feedback were obtained for 13 of the 14 participants. All 13 participants thought it would be helpful for the sensor to vibrate when placed in the correct spot on the neck (mean 1.85, SD 0.38) as well as when swallowing exercises were performed correctly (mean 2.0, SD 0.00; Table 4). A total of 11 (85%) participants felt it would be helpful for the sensor to give haptic feedback of swallow quality during at-home testing (mean 1.5, SD 0.88; Table 4).

Discussion
Principal Findings
To our knowledge, this is the first study to assess evaluations from patients with head and neck cancer of a wearable throat sensor in clinical settings with separate cohorts at varying time points along their treatment trajectory. Across all studies, the overall willingness to wear the sensor for 9 months during the first year after radiation was high and the perceived need was rated highly. However, study 1’s results should be interpreted with caution since the participation rate was 59%, with non-Hispanic and higher income/education patients more likely to complete the questionnaire. While study 2 and 3 used convenience samples for user testing, accrual rates were high (88%), even for those undergoing active treatment at the time of approach.

Direct comparison of our results with other works is not possible since the vast majority of published data regarding wearable devices equipped with mechanical, optical, biochemical, electrical, or acoustic sensors are pilot studies conducted with graduate students in a laboratory under highly controlled conditions [60-64]. While it did not test actual user engagement over repeated time points, it did gather patients’ opinions about the likelihood that they would wear the sensor for a period of several months. This question was asked in study 1 for patients who were only exposed to a photo of the proposed sensor, whereas patients and survivors in study 2 were asked this question after wearing the actual sensor while swallowing boluses of varying textures in a controlled setting. When searching for comparable studies that address extended user engagement with health technologies, the extant literature is limited to nonsensor research with mobile websites or apps [65] and to real-world studies of fitness tracker abandonment rates in healthy adults; these studies tend to describe a steep decline in user engagement over time. It is possible that our high rates of expressed willingness to wear the sensor for 9 months is due to the perceived usefulness of this device for this highly specialized problem.

Since the majority of participants (137/166, 83%) expressed willingness to wear the sensor for 9 months, data from those participants who were unwilling provided valuable insight into the potential barriers to its long-term use. Across all three studies, nearly 86% (24/28) of the unwilling participants perceived the 9-month testing period as too long. The second-most prevalent reason, that the sensor’s appearance would provoke unwanted attention, was endorsed by 56% (15/27) of the unwilling participants. The third-most frequent
reason was an unwanted increase in daily responsibilities (16/26, 62%). This was also borne out by spontaneous comments in study 3, when nearly all 14 patients communicated a preference for a more streamlined one-step application process, rather than the separate applications for the strain sensor and surface EMG electrodes. On the other hand, several of the unwilling participants were much more willing to wear the sensor for 9 months if the sensor could provide individual dysphagic risk feedback and were made more unobtrusive in appearance (Figure 3). These findings are consistent with other mHealth reports citing multiple aspects of participant burden [48] and social implications of the technology’s appearance [66] as being relevant constructs to user engagement.

**Bidirectional Communication**

Our data confirmed two other persuasive design principles: the desire for bidirectional communication (dialog support) with a team of clinical experts (system credibility). In all three studies, a large proportion of patients endorsed the rationale for the sensor (study 1: 115/138, 83.5%; study 2: 10/14, 71.4%; study 3: 12/14, 85.7%; i.e., that sensor data be processed and sent back with contextual explanations of their risk of dysphagia development). Furthermore, of the three proposed persuasive design features, feedback about dysphagia risk had the greatest impact in increasing willingness among all participants (Table 4). These findings point to the importance of fostering a sense of connectedness and reassurance between the user and the technology so that patients’ association between their own health behaviors and subsequent health outcomes can be continually reinforced [42]. Future plans for implementation include data visualization of near-time individualized risk for dysphagia in the form of an app that can be linked with the throat sensor. When asked about direct haptic communication with the sensor itself, patients in study 3 rated haptics as helpful, especially when unsure about correct placement on the throat and whether preventive exercises were being done correctly (Table 4). One patient commented that he was never really sure if he was performing the exercises correctly at home and was “just winging it.”

**Sensor and Adherence to Exercises**

The majority of participants (97/119, 82%) agreed that the sensor would serve as a reminder for them to do their speech pathology swallowing exercises. While the main goal of the sensor is to provide earlier detection of radiation-associated dysphagia, reminding patients to complete their swallowing exercises at home to counteract the development of dysphagia could be an additional benefit to this developing technology. Since personalized risk information is generally not sufficient in itself to increase exercise adherence per se [67], further user-centered testing would be needed to assess preferred modes of sensor feedback (e.g., within an app or coupled with virtual coaching) [68].

**Limitations**

Our study was conducted solely with survivors and patients attending clinical visits at MD Anderson, which generally requires high-quality insurance for access. Generalizability of our results are further limited by examining the demographic patterns among respondents versus nonrespondents in study 1. A total of 38% (21/55) of the eligible survivors did not complete the questionnaire despite repeated contact by the study team; nonresponders were significantly more likely to be Hispanic (P=.003), without a bachelor’s degree (P=.02), and of lower annual household income compared to respondents (P=.007). This is consistent with Rising et al’s [69] recent analysis of National Cancer Institute’s 2018 Health Information National Trends (HINTS) population survey data showing that nonusers of personal mHealth technologies were more likely to be older than 65 years and have lower incomes. Given the challenge of sustaining patient engagement in mHealth technology, future research should target these patients who fit within the above demographic profiles. Finally, the sample sizes for study 2’s and 3’s on-patient testing were constrained by the need to complete all testing within 1-week periods, as the sensors were applied/tested by visiting engineers and not MD Anderson research staff. It is conceivable that larger sample sizes might have produced a wider variation in response to the sensor’s features and perceived usefulness.

**Conclusion**

Large proportions of non-Hispanic well-educated patients with high-quality insurance and above-average incomes were receptive to the idea of wearing a personalized risk sensor for an extended period during the first year after radiation. User ratings of discomfort and difficulty were minimal; however, a significant minority of patients expressed concern with various aspects of the sensor’s burden and its appearance.

**Acknowledgments**

We would like to acknowledge our patients who participated in the study and Evalyne W Kamunyo for her assistance with study 1.

The primary funding for this study was provided by the National Institute of Dental Craniofacial Research (NIDCR DE019141). DJ Lipomi acknowledges support from the National Science Foundation (grant CBET-2223566). Support was provided, in part, by the Assessment, Intervention and Measurement Shared Resource through a Cancer Center Support Grant (CA16672; principal investigator: P Pisters, MD Anderson Cancer Center), from the National Cancer Institute, from the National Institutes of Health, and through the Duncan Family Institute for Cancer Prevention and Risk Assessment.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No 898571, awarded to RB.
Data Availability
The data that support the findings of this study are available from the corresponding author EHS upon reasonable request.

Conflicts of Interest
None declared.

References


Abbreviations

EMG: electromyography
HINTS: Health Information National Trends
IMRT: intensity-modulated radiation therapy
MBS: modified barium swallow
mHealth: mobile health

©Eileen H Shinn, Adam S Garden, Susan K Peterson, Dylan J Leupi, Minxing Chen, Rachel Blau, Laura Becerra, Tarek Rafeedi, Julian Ramirez, Daniel Rodriguez, Finley VanFossen, Sydney Zehner, Patrick P Mercier, Joseph Wang, Kate Hutcheson, Ehab Hanna, Darren J Lipomi. Originally published in JMIR Cancer (https://cancer.jmir.org), 28.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Co-Design, Development, and Evaluation of a Mobile Solution to Improve Medication Adherence in Cancer: Design Science Research Approach

Thu Ha Dang1,2,3, MIPH, MD; Nilmini Wickramasinghe4,5,6,7, PhD, MBA; Abdur Rahim Mohammad Forkan8, PhD; Prem Prakash Jayaraman9, PhD; Kate Burbury10,11, PhD; Clare O’Callaghan12,13, PhD; Ashley Whitechurch14, MANP-NP; Penelope Schofield1,2,7,11, PhD

1Department of Psychological Sciences, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia
2Digital Cancer Care Innovation, Department of Health Services Research, Peter MacCallum Cancer Centre, Melbourne, Australia
3Digital Health Cooperative Research Centre, Sydney, Australia
4Department of Health and Bio Statistics, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia
5Epworth Healthcare, Melbourne, Australia
6Optus Chair Digital Health, La Trobe University, Melbourne, Australia
7Iverson Health Innovation Research Institute, Swinburne University of Technology, Melbourne, Australia
8Digital Innovation Lab, Department of Computer Science and Software Engineering, School Software and Electrical Engineering, Swinburne University of Technology, Hawthorn, Australia
9Factory of the Future and Digital Innovation Lab, School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Melbourne, Australia
10Digital and Healthcare Innovation, Peter McCallum Cancer Centre, Melbourne, Australia
11Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Australia
12Caritas Christi and Psychosocial Cancer Care, St Vincent's Hospital, Melbourne, Australia
13Department of Medicine, St Vincent's Hospital, The University of Melbourne, Melbourne, Australia
14Department of Clinical Haematology, Peter MacCallum Cancer Centre & Royal Melbourne Hospital, Melbourne, Australia

Corresponding Author:
Thu Ha Dang, MIPH, MD
Department of Psychological Sciences
Swinburne University of Technology
John Street
Melbourne, 3122
Australia
Phone: 61 422703347
Email: thuhadang@swin.edu.au

Abstract

Background: Medication nonadherence negatively impacts the health outcomes of people with cancer as well as health care costs. Digital technologies present opportunities to address this health issue. However, there is limited evidence on how to develop digital interventions that meet the needs of people with cancer, are perceived as useful, and are potentially effective in improving medication adherence.

Objective: The objective of this study was to co-design, develop, and preliminarily evaluate an innovative mobile health solution called Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) to improve medication adherence among people with cancer.

Methods: Using the 4 cycles and 6 processes of design science research methodology, we co-designed and developed a medication adherence solution for people with cancer. First, we conducted a literature review on medication adherence in cancer and a systematic review of current interventions to address this issue. Behavioral science research was used to conceptualize the design features of SAMSON. Second, we conducted 2 design phases: prototype design and final feature design. Last, we conducted a mixed methods study on patients with hematological cancer over 6 weeks to evaluate the mobile solution.
Results: The developed mobile solution, consisting of a mobile app, a web portal, and a cloud-based database, includes 5 modules: medication reminder and acknowledgment, symptom assessment and management, reinforcement, patient profile, and reporting. The quantitative study (n=30) showed that SAMSON was easy to use (21/27, 78%). The app was engaging (18/27, 67%), informative, increased user interactions, and well organized (19/27, 70%). Most of the participants (21/27, 78%) commented that SAMSON’s activities could help to improve their adherence to cancer treatments, and more than half of them (17/27, 63%) would recommend the app to their peers. The qualitative study (n=25) revealed that SAMSON was perceived as helpful in terms of reminding, supporting, and informing patients. Possible barriers to using SAMSON include the app glitches and users’ technical inexperience. Further needs to refine the solution were also identified. Technical improvements and design enhancements will be incorporated into the subsequent iteration.

Conclusions: This study demonstrates the successful application of behavioral science research and design science research methodology to design and develop a mobile solution for patients with cancer to be more adherent. The study also highlights the importance of applying rigorous methodologies in developing effective and patient-centered digital intervention solutions.

(JMIR Cancer 2024;10:e46979) doi:10.2196/46979

KEYWORDS
cancer; behavioral science; design science research; digital; medication adherence; mobile solution; Safety and Adherence to Medication and Self-Care Advice in Oncology; SAMSON; mobile phone

Introduction

Background

Optimal adherence to medication is increasingly one of the top priorities in oncology care [1-3]. Medication adherence (MA) is “the extent to which patients take their medication as recommended by their health care provider” [4]. Despite this importance, the MA rate is very low: only 14% for some cancer regimens [3,5,6]. Poor MA negatively impacts the health outcomes of the patient [3,7-9] and increases pressure on health services and health care fiscal restraints [9,10].

MA is a complex and multifaceted phenomenon that can be influenced by 5 interacting dimensions: socioeconomic and health system factors as well as condition-, therapy-, and patient-related factors [11]. Patient-related factors are the most important [12] because adherence interventions may potentially make the most impact on these factors without necessarily having systemic solutions [11]. Multicomponent interventions that involve collective adherence strategies and are tailored to patients are likely more effective than single-strategy interventions in addressing these factors and improving adherence to oral anticancer medicines [13]. Technology can help to deliver multicomponent interventions more effectively and efficiently [13-15] without requiring too many extra resources, which are already scarce, from the health system [16].

With the rapidly evolving nature and increased uptake of information and communications technologies in the last 20 years [14,17], mobile phone–based interventions have been widely used to address the problem of medication nonadherence, specifically in cancer [18,19]. Literature reviews showed the potential of using technologies such as mobile solutions in promoting MA by providing patients rapid, continuous, and easy access to educational resources and symptom or side effect self-management strategies as well as facilitating direct patient-provider communication [11,15,17]. However, there is very limited evidence on how to develop mobile solutions that meet the needs of people with cancer, are perceived as useful, and are potentially effective in improving MA [1,13,19,20].

Research Context

In our previous research, we developed REMIND, which is a mobile health system to increase adherence to oral medication in people with chronic myeloid leukemia (CML) [21]. It comprises daily SMS text messages to provide drug reminders and symptom self-care advice, as well as nurse telephone consultations to promote adherence [21]. The development of REMIND was guided by the framework for the development of complex interventions [22]. To understand patients’ experiences of CML and identify their possible facilitators and barriers to adherence, a prior qualitative study was conducted [23]. To increase the acceptability of the intervention, stakeholders (eg, consumers and oncology professionals), were involved in reviewing iterative REMIND revisions and resource manuals [21]. The intervention content and delivery mechanisms were based on theories and available evidence [21].

Findings from the REMIND pilot study [21] showed that most patients reported episodes of nonadherence during the study period. Some reasons for their nonadherence were intentional [24], such as to reduce dose-dependent side effects. Some patients reported unintentional nonadherence [25] due to forgetfulness and miscommunication with health care providers [23]. Health care professionals (HCPs) had challenges in accurately assessing patients’ adherence status and identifying causes of nonadherence [23]. Users found REMIND generally acceptable to use and appreciated its benefits in establishing medication routines, resolution of symptom uncertainty, increased awareness of self-care, and informed decision-making [21].

REMIND had limitations. First, using a design framework specifically for digital interventions is crucial; yet, this was missing in the REMIND system’s development. Second, although stakeholder involvement was reported in the intervention’s development process, a genuine co-design process, defined as “meaningful end-user engagement in research design and includes instances of engagement that occur
across all stages of the research process and range in intensity from relatively passive to highly active and involved” [26], was not adopted. Third and last, patients reported some functional errors and expressed their need for an advanced iteration with improved functionality and presentation as well as a more user-friendly application [21].

Given the importance of the medication nonadherence problem that has not been well addressed and the gap in literature on how to develop acceptable, useful, and potentially effective digital interventions to solve the problem, as well as the need to resolve the identified limitations of REMIND, we combined design science research methodology (DSRM) and co-design to develop its new version, named Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) mobile health solution.

DSRM Cycles and Stages

Over the last couple of decades, design science research (DSR) [27] has been one of the main paradigms characterizing most information system research that aims to design and implement innovative technologies [28,29] through 3 design cycles: rigor, design, and relevance [30]. In 2007, Peffers et al [31] presented 6 process stages of the DSRM: problem identification and motivation, definition of the objective of the solution, design and development, demonstration, evaluation, and communication. Later, Drechsler and Hevner [32] extended the original DSRM with a fourth cycle (change and impact) to capture the dynamic nature of information system artifact design for volatile environments. Furthermore, the DSRM has been used in different health care contexts [29,33,34], demonstrating its importance in developing patient-centered digital health solutions. We adapted these 4-cycle and 6-process DSRM models to direct the steps required for the design and development of the SAMSON mobile health solution (hereinafter SAMSON) to improve MA in cancer. We present SAMSON and describe in detail the process of applying DSRM to design and develop it to answer the research question “How can DSRM be applied to enhance the initial mobile health system to provide a better user experience to improve MA to oral anticancer agents in adults with cancer?” Our study aimed to co-design, develop, and preliminarily evaluate SAMSON.

Methods

Overview

In this section, we explain how the 4 cycles and 6 processes of DSRM were adapted to design and develop SAMSON. Figure 1 presents the 4 DSRM cycles used to design and develop SAMSON. Table 1 illustrates how the 6-process DSRM models were applied in this study.

Figure 1. The adapted 4-cycle design science research methodology of the Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) mobile health solution. DSR: design science research; KB: knowledge base.
Table 1. Adapted 6-process design science research methodology (DSRM) applied to design and develop the Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) mobile health solution.

<table>
<thead>
<tr>
<th>DSRM stages</th>
<th>Interaction of DSRM cycles and stages</th>
<th>Approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1: problem identification and motivation</td>
<td>Cycle 1: change and impact impacts on stage 1</td>
<td>• Review literature on MA&lt;sup&gt;a&lt;/sup&gt; problems in cancer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Review literature on current MA interventions and their effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Identify problems in the current design</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Define a set of requirements in the new design</td>
</tr>
<tr>
<td>Stage 2: definition of the objective of the solution</td>
<td>Cycle 1: change and impact and cycle 2: relevance impact on stage 2</td>
<td>• Review literature on BSR&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Adapt BSR principles in design</td>
</tr>
<tr>
<td>Stage 3: design and development</td>
<td>Cycle 2: relevance, cycle 3: design, and cycle 4: rigor impact on stage 3</td>
<td>• Conceptualize design requirements and features</td>
</tr>
<tr>
<td>Stage 4: demonstration</td>
<td>Cycle 4: rigor impact on stage 4</td>
<td>• Test the design and acquire feedback from the design’s users</td>
</tr>
<tr>
<td>Stage 5: evaluation</td>
<td>Cycle 4: rigor impact on stage 5</td>
<td>• Evaluate the acceptability, usability, and potential effect of the intervention</td>
</tr>
<tr>
<td>Stage 6: communication</td>
<td>Stage 6 impact on cycle 3: design</td>
<td>• Report and publish the evaluation results</td>
</tr>
</tbody>
</table>

<sup>a</sup>MA: medication adherence.<br>
<sup>b</sup>BSR: behavioral science research.

**DSRM Cycles**

The change and impact cycle [32] ensures that SAMSON (internal environment) would fit for purpose in the context of the Australian health care system, cancer care, digital health, and patient environments (external environment). The internal environment here includes the designed mobile solution and the users (patients and oncology clinicians). This was defined through the process of problem identification and motivation (DSRM stage 1).

The relevance cycle links the key identified requirements of the users, including the users’ needs from REMIND’s pilot test, and the problems that they are facing in their environments. This was done through a range of discussions with SAMSON’s stakeholders and was demonstrated in DSRM stage 2 (definition of the objective of the solution) and impacted to stage 3 (design and development).

The co-design cycle (phase 1 and 2) consists of smaller cycles or phases (interacting iterative processes), including constructing the artifact, evaluating it, and using evaluation feedback to further refine it until a satisfactory design is achieved [27]. This cycle is the center of the research project because it is directed by the relevance cycle and the rigor cycle [33]. However, this is not a 1-way process because the results of the co-design cycle will then become a part of the relevance cycle. This cycle was performed in DSRM stage 3.

The rigor cycle links design science activities and grounded knowledge bases, such as the scientific theories, experience, and expertise that inform the DSRM project [33]. The scientific theories applied in this study include the health belief model (HBM) [35], self-determination theory (SDT) [36], and behavioral learning perspective (BLP) [11]. The impact of these knowledge bases on the SAMSON was demonstrated in DSRM stages 3 (design and development), 4 (demonstration), and 5 (evaluation). In parallel, the design and use of the SAMSON provide new knowledge (eg, the effect of this solution in terms of promoting adherence among people with cancer) to the external environment (Australian health care and cancer care context) in which the mobile solution is embedded. This process was rigorously validated in stage 5.

**DSRM Processes**

The SAMSON design comprises two phases: (1) prototype design and development and (2) final feature design and development (Figure 2).
Figure 2. Design and development phases of the Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) mobile solution.

Phase 1: Prototype Design and Development
The first phase (prototype design and development) started by defining a specific research problem (stage 1). The research problem focuses on MA problems in cancer, barriers to MA, and current MA interventions and their effects. To establish problem awareness, a literature review on MA in cancer and a systematic review of current MA interventions in cancer were conducted by THD and the research team (ARMF, NW, PPJ, and PS) to define MA barriers and facilitators as well as the characteristics of effective MA interventions [13]. Besides, feedback from HCPs and patients regarding REMIND was examined thoroughly by THD and PS to define necessary changes in the next iteration (SAMSON). Guided by the problem awareness, behavioral science research (BSR) was used by THD and PS to conceptualize preliminary design requirements (stage 2). Subsequently, the design requirements were translated into design features for the SAMSON prototype (SAMSON version 1), with ARMF leveraging the available features of the REMIND system in consultation with HD and the research team (stage 3). A test was then conducted on a convenient community sample to investigate whether the prototype works, examine its features, and propose more design requirements that may be helpful for patients (stages 4 and 5). Purposive and snowball sampling [37] were used to recruit participants to test SAMSON version 1. The convenient community sample included project team members, HCPs, and people in the community.

Phase 2: Final Feature Design and Development
Feedback from the testing of SAMSON version 1 initiated the second phase (final feature design and development; stage 3). In this phase, the problems detected in phase 1 were fixed. On the basis of feedback from participants in the SAMSON version 1 testing, THD and the research team returned to the literature and consulted the BSR to address the participants’ suggestions and develop the final designed features in SAMSON version 2. A preliminary evaluation of SAMSON version 2 was conducted among people with cancer (stages 4 and 5). Details of the preliminary evaluation study (hereinafter SAMSON evaluation) are presented in the following subsections. The results of the design and development of the SAMSON will be presented in publications (stage 6).

SAMSON Evaluation Methods

Study Design and Setting
This is a study with an explanatory sequential mixed methods design: a quantitative survey was conducted first, followed by qualitative interviews [38]. The quantitative study was conducted using a purpose-built questionnaire. The qualitative component consisted of in-depth interviews with a subset of patient participants. The mixed methods study design was applied to use the qualitative interviews to explore and make sense of the quantitative findings [38].

Participants
Purposive sampling [39] was used to select outpatient patients from the hematology department at Peter MacCallum Cancer Centre in Melbourne, Victoria, Australia. To be eligible, participants were required to be adults (aged ≥18 years); have an established diagnosis of chronic lymphocytic leukemia, CML, essential thrombocythemia, malignant neoplasm, myelofibrosis, myeloproliferative neoplasms, or polycythemia rubra vera; be taking or commencing an oral anticancer medication; and have smartphone and internet access. Before participating in the study, the study staff helped participants to install the SAMSON app on their mobile phone and briefed them on how to use it. They also received the SAMSON app user manual with detailed information, including app introduction, features, how to install and navigate, and common issues and how to solve them.
**Intervention**

The SAMSON has two elements: (1) a smartphone-based app to remotely prompt MA, monitor the patient’s side effects, and provide self-care advice; and (2) a web-based application to program the patient’s daily drug reminders and side effect surveys, and provide relevant drug information. In this evaluation study, patients were asked to trial the SAMSON smartphone app (the first element). The SAMSON web page (the second element) was used to populate daily drug reminders, weekly side effect surveys, and relevant patient information. Data collected on patients’ self-reported MA and side effects were uploaded and archived on the SAMSON web page.

**Measures**

Patients used the SAMSON app for at least 6 weeks. Subsequently, they were asked to complete the questionnaire via a Qualtrics (Qualtrics International Inc) link [40] that they received in an email from a researcher. The purpose-built questionnaire was adapted from the Evaluation Tool for Mobile and Web-Based eHealth Interventions (Enlight) [41]. The items in the questionnaire were language adapted for respondents without a background in IT and health. Next, face validity testing [42] was applied to achieve a consensus on the adapted Enlight questions. Finally, usability testing following the think-aloud protocol [43] was conducted on 2 consumers to finalize the questionnaire for use.

The questionnaire assesses the quality of the SAMSON app on 6 main constructs or dimensions: usability, visual design, user engagement, content, therapeutic persuasiveness, and general evaluation. Each dimension had between 3 and 6 items, for a total of 25 items. The stem of the item was presented as a statement (eg, “Overall, I found the mobile app was easy to use”), and the response scale was a 5-point Likert scale (1=strongly disagree, 2=disagree, 3=neither, 4=agree, and 5=strongly agree; Multimedia Appendix 1).

**Qualitative Interview**

All interviews were conducted either face-to-face or via web-based videoconference platform (Zoom; Zoom Video Communications, Inc) by THD, using a semistructured interview guide [39] (Multimedia Appendix 2). Each interview lasted between 45 and 90 minutes and was audio recorded.

**Data Analysis**

**Quantitative Data**

Descriptive statistics were used to analyze the quantitative data. SPSS (version 28.0; IBM Corp) [44] was used.

**Qualitative Data**

Each interview was transcribed verbatim [39]. NVivo 12 qualitative data management software (Lumivero) [45,46] was used. Qualitative data were analyzed thematically using a comparative, iterative, and predominantly inductive process, informed by grounded theory [47,48]. Thematic analysis has been used widely in information system research for different purposes, such as to understand phenomena related to information systems [49] or to evaluate the effectiveness of IT artifacts [50]. A qualitative interrating process was also conducted [51]. First, THD completed coding all interview records. Next, CO reviewed all interviews as well as THD’s codes and agreed or disagreed with each code and also suggested additional codes. Subsequently, both researchers discussed the codes until they reached agreement. Codes were then collated into subcategories (labels for comparable code groups), categories (labels for comparable subcategory groups), and themes (labels for comparable category groups). THD led category and thematic development, which was followed by a review of the categories and themes by CO. All disagreements were also discussed, and adjustments were made until consensus was reached. Both authors reviewed the data to ensure that the themes worked in relation to the entire data set and to generate a thematic map of the analysis. The researchers’ interrating process helped to strengthen the credibility and trustworthiness of this study [52,53].

**Ethical Considerations**

The SAMSON evaluation was approved by the human research ethics committees of Peter MacCallum Cancer Centre (HREC/74134/PMCC) and Swinburne University of Technology (20215811-8152). Written consent was obtained from all participants. Throughout the comprehensive consent process, participants were informed that their participation in this research was voluntary and that they were free to withdraw at any stage if they wished to do so. In addition, they were informed that their data would be deidentified for analysis and publication. Participants did not receive any compensation from the research team.

**Results**

The results are presented in the sequence of DSRM stages as shown in Figure 2: review literature, review and adapt BSR, co-design and test SAMSON version 1 (design cycle 1), develop SAMSON version 2, and SAMSON evaluation (design cycle 2).

**Review Literature**

The literature review of most recent research on MA in cancer showed that the problem of medication nonadherence in cancer is still persistent [5,15]. The results of the systematic review of intervention solutions to enhance adherence to oral anticancer medicines in adults [13] were in line with those of earlier reviews of the same topic [54,55]: multidimensional interventions that use collective strategies (educational, reminder, cognitive, behavioral, and affective) to promote adherence were potentially more effective than single-strategy interventions. This could be explained because MA is a complex and multifaceted phenomenon determined by 5 dimensions—socioeconomic and health system factors as well as condition-, therapy-, and patient-related factors—that require different strategies to address [11,13]. The review also suggested that a combination of cognitive and behavioral theories may better explain the diverse barriers and facilitators to MA and provide stronger direction to formulate interventions [13].

**Review and Adapt BSR**

Guided by the problem awareness from the literature review and REMIND studies, we conducted a review of BSR to select...
cognitive and behavioral theories that can potentially address MA barriers and promote MA facilitators via the SAMSON. The HBM [35], SDT [36], and BLP [11] were chosen to govern the design requirements of the SAMSON mobile solution [56]. According to the HBM, people will take health actions (e.g., adherence) if they have 4 basic conditional beliefs or perceptions regarding the disease, the effect of the disease on people’s lives, the action to respond to the disease, and the conviction that the benefit of action will outweigh the barriers [57]. According to the SDT, motivation (intrinsic and extrinsic) is crucial to successful behavior change [36]. The behavioral theory emphasizes the role of positive and negative reinforcements in controlling people’s behaviors that are immediately relevant to adherence [11].

**Co-Design and Test SAMSON Version 1**

Using knowledge gained from the literature and core theories, as well as users’ feedback on REMIND’s limitations, preliminary design requirements were conceptualized. The outcome of such conceptualization is presented in Figure 3.

**Figure 3.** Conceptualization of Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) design requirements using wireframes.

During the SAMSON design stage, we focused on the following target behaviors: knowledge, reinforcement, intentions, emotion, social influences, beliefs about capabilities, behavioral regulation, memory, and attention. These behaviors were originated from key barriers to MA, considered most feasible to influence, and expected to contribute most to the improvement of adherence. On the basis of behavioral analysis of these behaviors, potential behavior change techniques (BCTs) from the HBM, SDT, and BLP as well as intervention functions were selected for the SAMSON app; for example, the *prompts or cues to action* technique from the HBM [58] was applied for the *medication reminder* feature. *Problem-solving* and *self-monitoring* techniques from the HBM and SDT [36] were applied for the *symptoms assessment and management* module of the app. The *feedback on behavior* technique from the BLP [11] was applied for the *reinforcements* module. **Textbox 1** shows the conceptual model picturing this process using BSR, including the selection of final BCTs and the app’s features.
**Textbox 1.** Conceptualizing Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) design requirements and features using behavioral science research.

<table>
<thead>
<tr>
<th>Requirements and features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Medication adherence barriers</td>
</tr>
<tr>
<td>• Drugs’ side effects</td>
</tr>
<tr>
<td>• Lack of medication knowledge</td>
</tr>
<tr>
<td>• Lack of motivation</td>
</tr>
<tr>
<td>• Lack of health care professional (HCP) support to manage side effects</td>
</tr>
<tr>
<td>• Poor patient-HCP communication</td>
</tr>
<tr>
<td>• Lack of self-efficacy</td>
</tr>
<tr>
<td>• Forgetfulness</td>
</tr>
<tr>
<td>2. What needs to change</td>
</tr>
<tr>
<td>• Knowledge</td>
</tr>
<tr>
<td>• Reinforcement</td>
</tr>
<tr>
<td>• Intentions</td>
</tr>
<tr>
<td>• Emotion</td>
</tr>
<tr>
<td>• Social influences</td>
</tr>
<tr>
<td>• Beliefs about capabilities</td>
</tr>
<tr>
<td>• Behavioral regulation</td>
</tr>
<tr>
<td>• Memory and attention</td>
</tr>
<tr>
<td>3. Behavior change techniques from the theories</td>
</tr>
<tr>
<td>• Information about side effects and medicines</td>
</tr>
<tr>
<td>• Feedback on behavior</td>
</tr>
<tr>
<td>• Social support</td>
</tr>
<tr>
<td>• Problem-solving</td>
</tr>
<tr>
<td>• Self-monitoring behavior</td>
</tr>
<tr>
<td>• Prompts</td>
</tr>
<tr>
<td>• Habit formation</td>
</tr>
<tr>
<td>4. App features</td>
</tr>
<tr>
<td>• Side effects section</td>
</tr>
<tr>
<td>• Medication information section</td>
</tr>
<tr>
<td>• HCPs’ contacts</td>
</tr>
<tr>
<td>• HCP connections</td>
</tr>
<tr>
<td>• Side effects self-management section</td>
</tr>
<tr>
<td>• Drug reminders</td>
</tr>
<tr>
<td>• Reinforcements</td>
</tr>
</tbody>
</table>

The aforementioned step is followed by a translation into design features for prototype implementation. The features were designed to provide a solution within 1 IT artifact, which is called the SAMSON mobile solution, including a mobile app, a web portal, and a cloud-based database for storing patient-specific information. The mobile app is available for patients to use, while the web portal is available for both patients and their care team. The SAMSON included 5 different modules with some key requirements (Textbox 2).
The 5 different modules of the Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON).

<table>
<thead>
<tr>
<th>Medication reminder and acknowledgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The app should support and display multiple medications and send a reminder per medication. The reminder can address the medication adherence (MA) barrier of forgetfulness. Patients need to tap each reminder for an acknowledgment.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symptom assessment and self-care management</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A list of available side effects and symptoms and self-care management in the mobile app is displayed for the medication that the patient is taking. Patients should be able to complete a symptom assessment survey through the app that will be distributed to patients using app reminders. They should be able to view information on how to manage their symptoms (if minor) and when they need to contact health care professionals (HCPs). This provides patients with medication knowledge as well as support in side effects self-management, both of which are important enablers of MA.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reinforcement</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The app sends a positive reinforcement to the patient at a specific time each week based on the MA profile for that week. Positive reinforcement can help motivate patients’ adherence.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Patients can use the app to view their profile information, such as their basic personal and clinical information, emergency contact and care team contact details, and medication information (both general and important). This information can address the MA barriers of lacking or misunderstanding medication information and poor patient-HCP communication.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Analytical reports of patients’ adherence status and their symptoms are available on the web portal for HCPs and patients in real-time. HCPs should be able to use these data in monitoring patients and providing them in-time and tailored support to manage side effects as well as to overcome MA barriers.</td>
</tr>
</tbody>
</table>

Test SAMSON Version 1
SAMSON version 1 was tested by 21 participants from a convenient community sample, which included project team members, HCPs, and people in the community. We sought participants’ feedback on issues regarding the expected features and functionalities of the prototype and its visual design. Overall, participants reported some functional errors, such as misdelivered medication reminders and data entry failures in some fields both in the smartphone app and on the web page. They also asked for new visual design requirements to meet users’ needs (details are presented in Multimedia Appendix 3).

Develop SAMSON Version 2
Overview
In phase 2, consumers’ feedback from design cycle 1 was reviewed by the project team. We grouped them based on the artifact’s functions and priority in terms of improvement (Multimedia Appendix 3). The main improvements in SAMSON version 2 are described in terms of priority in the following subsections.

Priority 1: Fix Functional Errors
All functional errors were fixed in this stage, including misdelivered medication reminders, app log-in-related issues, slow responsiveness to load app content, editing errors of medication schedules on the website, and functional errors related to data saving and data sorting on the website.

Priority 2: Enhance Existing Features and Functions
Textbox 3 presents feature and function enhancements in SAMSON version 2 in comparison to version 1.
Textbox 3. Enhanced features and functions in Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) version 2 in comparison to version 1.

<table>
<thead>
<tr>
<th>Version 1</th>
<th>Version 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The medication reminders had no expiry time</td>
<td>1. Setting up an expiry time (6 hours) for the reminders</td>
</tr>
<tr>
<td>2. Medications did not have color attributes</td>
<td>2. Adding color attributes for medications on the website</td>
</tr>
<tr>
<td>3. Patients could not view their data on the website</td>
<td>3. Enabling patients to log in to the website to view their own adherence performance, symptom reports, and completed side effects surveys</td>
</tr>
<tr>
<td>4. Patients could not export side effects surveys from the website to their data folder</td>
<td>4. Enabling patients to export the side effects surveys from the website to Excel (Microsoft Corp) spreadsheets</td>
</tr>
</tbody>
</table>

Priority 3: Create New Functions

The need expressed by consumers for new functions was discussed by the project team and conceptualized to guide the development of new selected design features in the new iteration of SAMSON. The literature, BSR theories, and BCTs were revisited when necessary to address new suggested requirements. Some selected new functions of SAMSON version 2 in comparison to version 1 are described in Multimedia Appendix 4.

SAMSON Evaluation

Overview

In the SAMSON evaluation, 30 (81%) of the 37 patients who were approached agreed to participate in the study and used SAMSON. None of them withdrew from the study. After 6 weeks, of the 30 participants, 27 (90%) completed the questionnaire, and 25 (83%) participated in the interview (Multimedia Appendix 5). Data retrieved from the SAMSON web page showed that 100% (1890/1890) of the drug reminders were sent on time, and all participants responded to the reminders and viewed the reinforcement messages. Most of the participants (23/30, 77%) reported side effects during the study period.

Participant demographics are described in Table 2. The mean age of the patient participants was 57.6 (SD 12.5) years. Most of the participants (20/27, 74%) were male individuals. The average time that participants had received treatments before the start of the study was 7.2 (SD 6.7) years. Approximately two-thirds of the participants (17/27, 63%) lived in the metropolitan areas of Melbourne. The proportions of participants using iPhones and Android mobile phones were equal.
Table 2. Participant demographics and clinical characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>57.6 (12.5)</td>
</tr>
<tr>
<td><strong>Sex (n=27), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20 (74)</td>
</tr>
<tr>
<td>Female</td>
<td>7 (26)</td>
</tr>
<tr>
<td><strong>Country of birth (n=27), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>14 (52)</td>
</tr>
<tr>
<td>Croatia</td>
<td>1 (4)</td>
</tr>
<tr>
<td>India</td>
<td>3 (11)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Pakistan</td>
<td>1 (4)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Not provided or missing</td>
<td>5 (19)</td>
</tr>
<tr>
<td><strong>English as first language (n=27), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25 (93)</td>
</tr>
<tr>
<td>Not provided or missing</td>
<td>2 (7)</td>
</tr>
<tr>
<td><strong>Education (highest level completed; n=27), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No formal schooling or incomplete schooling</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Primary school</td>
<td>9 (33)</td>
</tr>
<tr>
<td>Secondary or high school</td>
<td>3 (11)</td>
</tr>
<tr>
<td>Vocational</td>
<td>6 (22)</td>
</tr>
<tr>
<td>University</td>
<td>7 (26)</td>
</tr>
<tr>
<td>Postgraduate diploma or master’s degree or PhD</td>
<td>1 (4)</td>
</tr>
<tr>
<td><strong>Employment status (n=27), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Working full-time</td>
<td>8 (30)</td>
</tr>
<tr>
<td>Working part-time</td>
<td>5 (19)</td>
</tr>
<tr>
<td>Casual</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Sick leave (permanent)</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Retired</td>
<td>9 (33)</td>
</tr>
<tr>
<td><strong>Residence (n=27), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Metropolitan</td>
<td>17 (63)</td>
</tr>
<tr>
<td>Rural</td>
<td>10 (37)</td>
</tr>
<tr>
<td>Time since diagnosis (years), mean (SD)</td>
<td>7.2 (6.7)</td>
</tr>
<tr>
<td><strong>Diagnosis (n=30), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Chronic lymphocytic leukemia</td>
<td>15 (50)</td>
</tr>
<tr>
<td>Chronic myeloid leukemia</td>
<td>12 (40)</td>
</tr>
<tr>
<td>Essential thrombocytethemia</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Myeloproliferative neoplasms</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Polycythemia rubra vera</td>
<td>1 (3)</td>
</tr>
<tr>
<td><strong>Mobile phone operating system (n=30), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Android</td>
<td>15 (50)</td>
</tr>
<tr>
<td>iOS</td>
<td>15 (50)</td>
</tr>
</tbody>
</table>
Quantitative Survey

Participants’ responses to the Enlight questionnaire are presented in Multimedia Appendix 6. Enlight aims to examine individual quality constructs or dimensions, which means it is a suite of scales rather than 1 quality measure; therefore, we did not present the overall scale of the questionnaire.

Usability assesses the ease of learning how to use an app and the ease of using it properly. Overall, of the 27 participants, 21 (78%) rated the app as easy to use; only 1 (4%) participant rated it as difficult.

Visual design assesses the look and feel of an app. Participants mentioned that the app is attractive (14/27, 52%); well organized (19/27, 70%); and has appropriate font size, buttons, and menus (23/27, 85%). Some of them expressed a need for the app to be redesigned to increase its appeal (3/27, 11%) and encourage engagement (5/27, 19%).

User engagement assesses the extent to which an app’s design attracts people to use it. In general, participants were interested in using the app (18/27, 67%) because it was presented in an interesting way (19/27, 70%), different features were used to increase users’ interactions (19/27, 70%), automated features to respond to the survey were easy to use (21/27, 78%), and the app’s features were personalized to users (23/27, 85%). However, of the 27 participants, 2 (7%) were not interested in using the app at all.

In terms of the content, more than two-thirds of the participants (19/27, 70%) were satisfied with the amount of information and the way it was presented in the app. Nevertheless, 7% (2/27) of the patients thought that information about the app’s purpose was missing. Patients also reported that information about drugs and side effects was presented with gaps, overexplanation, or irrelevance (5/27, 19%).

Therapeutic persuasiveness assesses the extent to which an app is designed to encourage a patient’s MA. High proportions of participants agreed that the app provided activities to improve their adherence (21/27, 78%) and appropriate ongoing feedback (19/27, 70%). However, approximately one-fifth of the participants (6/27, 22%) did not think that completing activities on the app would help them to be more adherent to treatments. Patients thought that the app did not fully disclose information on how it can help them to be more adherent (7/27, 26%) and what they need to do for this (8/27, 30%).

Overall, 18 (67%) of the 27 participants thought that the app was valuable in assisting MA via improving their confidence in complying to treatments (11/27, 41%) and motivating them to do so (15/27, 56%). More than half of the participants (17/27, 63%) would recommend the SAMSON app to other people with cancer.

Qualitative Interview

Overview

Three common themes were generated from the interview data: (1) SAMSON is a generally helpful app that can remind, support, and inform; (2) possible barriers encompass app glitches and users’ technical inexperience; and (3) users desire customization, health care connections, and content refinement of SAMSON (Figure 4). A full presentation of themes, categories, and subcategories is presented in Multimedia Appendix 7. Further clarification of the themes is provided in the following subsections.
SAMSON Is a Generally Helpful App That Can Remind, Support, and Inform

Participants valued the SAMSON app’s features (eg, medication reminders and side effects information) because they are reliable, both in terms of content and functioning. As patients had to take >1 drug, different drug schedules and a busy life made it easy to forget taking pills on time. However, with the use of SAMSON, patients were reminded to take their medication on time:

[U]sually the reminder comes very close to the time...I do like that. [P6]

[T]he medication reminder is very prompt, and I have been, you know, taking my medication regularly; absolutely, I never missed out. [P11]

Although patients had access to different sources of information about the drug and its side effects, they found it easy to obtain the appropriate information that they needed in SAMSON:

[W]hen you go into the individual drugs and you’ve got the side effects, information is much easier to access from the app than it is if you go online. Or if you go into the drug information sheet, which is just, you know, overwhelming, it’s difficult. Particularly for somebody without a medical or paramedical background. But I thought the content was really well done. So you know I could find what I needed to know...it was a very good summary. [P13]

A participant and their carer trusted the app’s information because it was based on reputable research:

[O]n the Samson app, you can go through and actually know you can trust what’s in there...you guys
Patients could also benefit from information about drugs and side effects in the app. It supported them in managing unpleasant symptoms as consequences of cancer treatments that can be discouraging:

The side effects part of the screen was helpful to me in a simple fashion...support for people that are trying to manage their side effects. I thought it was good...excellent. [P14]

SAMSON could also encourage MA maintenance via reinforcement messages:

I liked the way it encouraged. You know, I was like, you know, try better this week, you know, do better. [P1]

It is encouraging the patient to always use the app and to take the medicine on time. [P12]

Overall, most of the participants thought that SAMSON is a helpful digital solution to promote MA:

It does its job, so it’s good...It’s a very neat app in this in the sense there’s no extra stuff. It’s just exactly just what it needs to do. That’s all, so yeah, it’s pretty good. [P7]

However, some of the participants mentioned that they did not need the app to comply with the treatment because they either had a well-established drug-taking routine for many years or had another medication reminder strategy:

I’ve got a container with my medications in, as soon as I get them, I write the dates on there, I know exactly, I don’t need my phone, I don’t need an app, I don’t need anything to remind me, I know, it’s a routine that I’ve done for too long...I’ve been on it over 14 years, so for me it’s a daily thing. [P24]

Nonetheless, these participants still praised the app as helpful for other patients, especially for those who are new to the treatment and like to use technology:

I can see when someone’s first-time treatment it’d be very very useful. [P9]

A newer person coming into their treatment, or a younger person that’s a bit more tech savvy, would probably prefer to use technology as a reminder. And you know that would be very handy. [P10]

Possible Barriers Encompass App Glitches and Users’ Technical Inexperience

Despite the benefits that SAMSON can bring to patients, it has some functional errors and drawbacks. These could annoy patients and make the app less effective:

I’ve got an Android [mobile phone] and I had to refresh the page many, many, many times. [P14]

The notification does come up, but it sits in the background on your phone, so it comes up separate from the app, as a notification. But it just sort of sits in the background. [P30]

Besides, as in the case of other advanced technologies, the use of SAMSON could be challenging for some people, especially older adults and people who are not technology savvy. A participant reflected on how others might view the app:

I think an app like that for my father who’s in his 80s, I couldn’t see it being used, he’d see it as a nuisance. [P9]

Users’ Desires Related to SAMSON Include Content and Feature Refinement, Customization, and Connections to Carers and HCPs

Patients expressed their desire for refinements in the new version of SAMSON to make it more appealing and capable of serving diverse needs of different users:

Maybe you can increase the size or to magnify for people. [P31]

[P]ossibly people might find something that gives them their compliance, or you know a color changes, [signifying] you’re on track, no you’re not. That may help them. [P7]

New features were also suggested to improve SAMSON’s effectiveness for both patients and their clinicians in disease management:

It might be handy on the app somewhere for the person using it to be able to make a note and say, put dates in “I’ve been in hospital” or “I’ve had broken, been in a car accident” or “I’ve got some bruising as a result of a fall” or something. [P10]

You probably need to have areas where people can actually add things to it, other than just keep going click click click and then get nothing at the end...it doesn’t really help...a section where you can add additions to it, even say a basic of when there’s a section on gastro and vomiting...did it affect you for a percentage...and then you may be able to assist from that side. [P24]

App modifications were also advised to improve patients’ proactiveness in treatment management:

It makes you feel more like you’re in control and that you can think you’re more likely to use an app if you can customize it to meet your own needs. And, whereas you know if I wanted to change it, and then I had to get in touch with someone to do that. Yeah, it’s just a bit disempowering. [P27]

Furthermore, patients emphasized new features to assist carer engagement with the app, which would support their MA, and communication with their HCPs when needing additional support:

Some patients, they don’t have this ability to manage their own medications, even when they have the app, and they need carer or family member to be also involved in the app. [P12]

There were times in the past that I might’ve had a side effect, or something had happened, and sometimes it was very hard to contact the nurse that’s
Discussion

Overview

This research study co-designs, develops, and evaluates an innovative mobile health solution to improve MA in cancer. Preliminary results demonstrate the successful application of BSR and DSRM to enhance the initial mobile health system and provide a better user experience. The study contributes to theory and practice in many ways.

Theoretical Contributions

Our study contributes to DSRM theory in 4 different ways. First, we expanded the scope of DSRM by applying it to the design and development of a mobile health solution for MA. Given the current challenges in public health and clinical fields, the potential of using DSRM to improve the effectiveness and efficiencies of health care innovations is enormous [33]. DSRM has been used to design new artifacts in different health care application areas [29,34] (eg, medical devices), but none of them target MA in cancer [13,33]. Hence, this study sheds new light on how DSRM can be applied in this area.

Second, we addressed the knowledge gap on how BSR and DSRM can be integrated to develop engaging and effective behavior change digital health solutions. There is a strong view that design science and behavioral science are 2 distinct research paradigms [27]. Design science is related to the creation of new artifacts, while behavioral science studies behavior in relation to IT use [59]. While behavioral science could be seen as a reactive and retrospective process to explain what already exists, design science is more proactive in its way in terms of creating technological solutions for the future [59]. Nevertheless, these 2 seemingly divergent research paradigms have some similarities. They both emphasize the importance of understanding the health problem before designing a solution and aim to ensure that the designed solutions can effectively engage users [60]. Engagement with mobile health interventions is a precondition for their effectiveness [61]. Behavior change theory and BCTs can assist macroengagement with the behavior changes the mobile health intervention aims to support (eg, MA) [62], while design science approaches, such as user-centered design, is more likely to facilitate microengagement with the mobile health interface (eg, logging in to the app) [60,63]. Therefore, integrating best practices from BSR and DSRM can bring more mutual benefits to design engaging behavior change interventions [60]. Research also showed that digital interventions developed using behavior change theory and BCTs are more likely to be effective than those without [60,62]. However, little is known about how these 2 approaches can be blended throughout the design process of artifacts to ensure that microengagement and macroengagement needs will be met [60]. Here, our study provided more understanding about how this integration can be done.

Third, an explanatory sequential mixed methods design [38] was applied in the evaluation stage of the SAMSON’s DSRM. This type of design is helpful for us to know why the user rated the solution’s quality as low or high for each criterion and gain further understanding on how we can improve the SAMSON in its next version. Because of the assumption that technical knowledge is needed to complete the questionnaire, we adapted Enlight for a lay audience. Unlike some other assessment tools, Enlight includes some quality constructs associated with intervention outcomes, such as persuasive design, behavior change, or therapeutic alliance [41], which is specifically necessary for a mobile solution, such as SAMSON, to change patients’ behavior toward medical treatments. This tool has been validated for assessing eHealth interventions regardless of delivery mediums and clinical aims [41]. In our study, it was language adapted but requires further validation for a community sample. Measures for evaluating the quality of a designed artifact are often difficult to define and are controversial [33]. Therefore, applying a mixed methods design with an appropriate, reliable, and valid assessment tool in the evaluation of digital interventions (eg, in the case of SAMSON) could be one of the effective ways to address this challenge.

Fourth, we effectively involved stakeholders, including real users, early and throughout the co-design and evaluation processes. We formed a project steering committee that included experts in allied health, app development, computer sciences, digital health, oncology, and psychology, as well as consumer representatives. They were involved very early in the co-design process to guide the review of behavior change literature and the selection of targeted change in nonadherence behavior. The committee was also involved in reviewing problem identification and design motivation, adapting BSR principles, and conceptualizing design requirements for SAMSON. After development, SAMSON was thoroughly tested by reasonable numbers of users (21 consumers tested version 1, and 30 patients tested version 2). Their feedback in the testing was then used to construct new requirements or refine the next version of SAMSON. By recognizing users as experts of their own experience, the proper co-design process can address pitfalls in the design and development of mobile health solutions that might limit adoption and effective use in practice [64-66] by facilitating necessary collaborations of diverse stakeholders [67,68] and leveraging expert insights and best practices [69,70].

Our study also contributes to the literature of interventions to promote MA in cancer. Systematic reviews of MA interventions in adults with cancer showed that there was limited use or poorly reported use of theory [71] and frameworks [20] in the design and development of digital interventions [13]. A high number of MA digital interventions have been proposed, but many of them have low user acceptance [72], and their effectiveness in clinical oncology practice is poorly supported [17,20]. Perhaps poor design is 1 reason for these issues [33]. To the best of our knowledge, SAMSON is one of the very first mobile health solutions to improve MA in cancer that applied rigorous DSRM in the design and development process. The use of DSRM provided various improvements in identifying and addressing requirements as well as evaluating this digital solution.
SAMSON was perceived as acceptable, usable, and useful by end users.

**Implications for Practice and Future Research**

Broadly, our study’s findings have implications for behavioral science and design science researchers, MA intervention developers, and oncology care providers. These findings provide additional evidence on the use of DSRM in health innovations. They can be used to develop principles for guiding DSRM adaptation and BSR integration in the design and development of mobile health solutions in general as well as those targeting MA. The findings of this work provide insights for oncology care providers to use, while encouraging the use of digital solutions to promote MA and drive health care outcomes. Technologies can enhance measures to improve MA, such as patient education as well as side effect monitoring and reporting, and facilitate effective communication between patients and their care team.

Our respondents indicated their acceptance of the mobile solution and valued its usability and usefulness in supporting their adherence to medication. They also reported some functional errors and the need for some further improvements in the design and features of SAMSON. We will use these findings to refine SAMSON and evaluate its acceptability, usability, and effectiveness in a future randomized controlled trial. On the basis of the feedback of participants, in the trial’s protocol, we will include assessments to help identify those who would benefit from the SAMSON.

**Limitations**

This study has limitations. Participants enrolled in the SAMSON preliminary evaluation are from the Peter MacCallum Cancer Centre hematology department, and most of them used only 1 oral anticancer regimen. As Peter MacCallum Cancer Centre is one of the leading oncology hospitals in Australia, in the interview, patients acknowledged that the care service that they received was of high quality. Many were provided medication education before starting treatments and at ongoing follow-up appointments. As a result, their perceptions of MA solutions may not represent those of patients who use multiple anticancer medicines and receive care from low-resource oncology settings. Future research can extend the evaluation of SAMSON to patients with other types of cancer at different levels of oncology care institutions.

**Conclusions**

By following the systematic DSRM approach, a patient-centered mobile health solution was developed to meet the needs and preferences of people with cancer and thus highly likely to be used by end users. This extensive report of the intervention development process provides transparent guidance on how to develop patient-centered digital mobile health solutions that will have a high likelihood of uptake.

**Acknowledgments**

The authors thank all patients who participated in the Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) evaluation study for their time and valuable feedback on the SAMSON. The authors thank Marliese Alexander and Senthil Lingaratnam for their review and guidance on the drug information aspect of the study. The authors also thank Jim Coomes and Hilary Schofield for their contribution in the development of the Evaluation Tool for Mobile and Web-Based eHealth Interventions (Enlight). The authors highly appreciate support from the Peter MacCallum Cancer Centre hematology department in the study participants’ recruitment. This publication is supported by the Digital Health Cooperative Research Centre (DHCRC), Swinburne University of Technology, and Peter MacCallum Cancer Centre (project DHCRC-0043). The DHCRC is funded under the Commonwealth’s Cooperative Research Centres program. THD is supported by the Australian Government Research Training Program Scholarship. The sponsor had no influence on the study design or the collection, analysis, and interpretation of data. The final decision to include the peer reviewers’ comments and submit the manuscript for publication was made only by the authors.

**Data Availability**

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

**Authors’ Contributions**

All authors have made substantial contributions and approved the conceptions, drafting, and final version of the manuscript. THD undertook the analysis and interpretation of the data. THD drafted the paper with contribution from NW, ARMF, PPJ, KB, CO, AW, and PS.

**Conflicts of Interest**

None declared.

Multimedia Appendix 1

Evaluating the Safety and Adherence to Medication and Self-Care Advice in Oncology (SAMSON) mobile app questionnaire. [DOC File, 76 KB - cancer_v10i1e46979_app1.doc ]
References


54. Unni E, Bae S. Exploring a new theoretical model to explain the behavior of medication adherence. Pharmacy (Basel) 2022 Apr 01;10(2):43 [FREE Full text] [doi: 10.3390/pharmacy10020043] [Medline: 3548702]


64. Noorbergen TJ, Adam MT, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021 Nov 10;9(11):e27896 [FREE Full text] [doi: 10.2196/27896] [Medline: 34757323]


Abbreviations

BCT: behavior change technique
BLP: behavioral learning perspective
BSR: behavioral science research
CML: chronic myeloid leukemia
DSR: design science research
DSRM: design science research methodology
Enlight: Evaluation Tool for Mobile and Web-Based eHealth Interventions
HBM: health belief model
HCP: health care professional
MA: medication adherence
SAMSON: Safety and Adherence to Medication and Self-Care Advice in Oncology
SDT: self-determination theory

Edited by T de Azevedo Cardoso; submitted 04.03.23; peer-reviewed by XF Zhong, J Seitz, E Aigbokhan; comments to author 10.01.24; revised version received 27.01.24; accepted 14.02.24; published 03.04.24.

Please cite as:
Co-Design, Development, and Evaluation of a Mobile Solution to Improve Medication Adherence in Cancer: Design Science Research Approach
JMIR Cancer 2024;10:e46979
URL: https://cancer.jmir.org/2024/1/e46979
doi:10.2196/46979
PMID:38569178
Telehealth With Comprehensive Live-Fed Real-World Data as a Patient Care Platform for Lung Cancer: Implementation and Evaluation Study

Di Zheng1*, MD; Yanhong Shang2*, MD; Jian Ni1, MD; Ling Peng3, MD; Xiaoming Tan4, MD; Zhaoxia Dai5, MD; Yizhuo Zhao6, MD; Aiqin Gu6, MD; Jiying Wang1, MD; Yanyan Song7, MD; Xiaofeng Li8, MD; Junping Zhang9, MD; Wei Heng10, MD; Cuiying Zhang11, MD; Chunling Liu12, MD; Hui Li11, MD; Yingying Du13, MD; Jianfang Xu1, MD; Dan Wu14, MD; Xuwei Cai6, MD; Rui Meng15, MD; Xiaorong Dong15, MD; Yaoping Ruan16,17*, PhD; Liyan Jiang6*, MD

1Department of Medical Oncology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China
2Department of Medical Oncology, Affiliated Hospital of Hebei University, Baoding, China
3Department of Respiratory Disease, Zhejiang Provincial People's Hospital, Affiliated People's Hospital, Hangzhou Medical College, Hangzhou, China
4Department of Respiratory and Critical Care Medicine, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
5Department of Thoracic Medical Oncology II, The Second Hospital of Dalian Medical University, Dalian, China
6Department of Respiratory and Critical Care Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
7Department of Biostatistics, Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, China
8Internal Medicine of Thoracic Oncology, Baotou Tumor Hospital, Baotou, China
9Department of Thoracic Oncology, The Affiliated Bethune Hospital of Shaanxi Medical University, Taiyuan, China
10Department of Medicine, Respiratory, Emergency and Intensive Care Medicine, The First Affiliated Hospital of Soochow University, Suzhou, China
11Cancer Center, Inner Mongolia Autonomous Region People's Hospital, Huhhot, China
12The Second Ward, Department of Pulmonary Medicine, Xinjiang Medical University Affiliated Tumor Hospital, Urumqi, China
13The First Affiliated Hospital of Anhui Medical University, Hefei, China
14Department of Thoracic Surgery, Cixi People's Hospital, Ningbo, China
15Cancer Center, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
16College of Mathematics and Computer Science, Zhejiang A&F University, Hangzhou, China
17Metaframe Technologies Inc, Shanghai, China
*these authors contributed equally

Corresponding Author:
Liyan Jiang, MD
Department of Respiratory and Critical Care Medicine
Shanghai Chest Hospital
Shanghai Jiao Tong University School of Medicine
No 241 West Huaihai Rd
Shanghai, 200030
China
Phone: 86 13916146759
Email: jiang_liyan2000@126.com

Abstract

Background: Telehealth has emerged as a popular channel for providing outpatient services in many countries. However, the majority of telehealth systems focus on operational functions and offer only a sectional patient journey at most. Experiences with incorporating longitudinal real-world medical record data into telehealth are valuable but have not been widely shared. The feasibility and usability of such a telehealth platform, with comprehensive, real-world data via a live feed, for cancer patient care are yet to be studied.
Objective: The primary purpose of this study is to understand the feasibility and usability of cancer patient care using a telehealth platform with longitudinal, real-world data via a live feed as a supplement to hospital electronic medical record systems specifically from physician’s perspective.

Methods: A telehealth platform was constructed and launched for both physicians and patients. Real-world data were collected and curated using a comprehensive data model. Physician activities on the platform were recorded as system logs and analyzed. In February 2023, a survey was conducted among the platform’s registered physicians to assess the specific areas of patient care and to quantify their before and after experiences, including the number of patients managed, time spent, dropout rate, visit rate, and follow-up data. Descriptive and inferential statistical analyses were performed on the data sets.

Results: Over a period of 15 months, 16,035 unique users (13,888 patients, 1,539 friends and family members, and 174 physician groups with 608 individuals) registered on the platform. More than 382,000 messages including text, reminders, and pictures were generated by physicians when communicating with patients. The survey was completed by 78 group leaders (45% of the 174 physician groups). Of the participants, 84% (556/678; SD 8.7) reported a positive experience, with efficient communication, remote supervision, quicker response to questions, adverse event prevention, more complete follow-up data, patient risk reduction, cross-organization collaboration, and a reduction in in-person visits. The majority of the participants (59/78, 76% to 76/78, 97.4%) estimated improvements in time spent, number of patients managed, the drop-off rate, and access to medical history, with the average ranging from 57% to 105%. When compared with prior platforms, responses from physicians indicated better experiences in terms of time spent, the drop-off rate, and medical history, while the number of patients managed did not significantly change.

Conclusions: This study suggests that a telehealth platform, equipped with comprehensive, real-world data via a live feed, is feasible and effective for cancer patient care. It enhances inpatient management by improving time efficiencies, reducing drop-off rates, and providing easy access to medical history. Moreover, it fosters a positive experience in physician-patient interactions.

(KEYWORDS: telehealth; real-world data; patient engagement; lung carcinoma; patient-reported outcomes)

Introduction

According to the GLOBOCAN 2020 report, the cancer mortality rate is higher in China than that in developed countries [1]. Lung cancer remains the most common and deadliest type of cancer, with an estimated 0.82 million new cases and 0.72 million deaths in 2020 in China [1]. In contrast to the rapidly declining mortality rate for lung cancer in high-income countries between 2000 and 2012 [2], the trend in the lung cancer mortality rate was stable in China from 2000 to 2016 [3]. Despite favorable survival outcome data for Chinese patients in international randomized clinical trials, these data do not reflect the real-world situation for the general population. The less-than-optimal progress in cancer control, especially in terms of the mortality rate, may be attributed to health care disparities between different regions, particularly urban and rural areas [4,5]. Clinical trial data from inadequately represented cancer patient populations could be complemented with real-world evidence to better inform health care practice and policy decisions [6].

The rapid development and adoption of new treatment regimens have made posttreatment care a critical factor in extending the cancer survival rate and improving patients’ quality of life [7]. Concurrently, telehealth has quickly become a major care delivery mechanism in recent years, a trend accelerated by the COVID-19 pandemic. One ongoing effort to sustain and scale digital health involves enabling data sharing and integration across different health systems [8]. Consequently, most telehealth systems today rely on point-in-time medical records that do not contain historical records nor data from other institutes. To overcome this data barrier, the platform implemented in this study has the capability to acquire medical records directly from patients.

Though there are perceivable benefits to having comprehensive medical records for telehealth, enabling comprehensive and longitudinal data for each patient involves tremendous effort. Such data are also critical for deriving conclusive real-world evidence [9]. Data acquisition must be inclusive, especially of vital signs related to the patient’s daily health status throughout the entire treatment period, in addition to diagnosis and treatment information [10]. Although this is currently achievable with the adoption of wearables and mobile devices, there are still tremendous challenges in longitudinally compiling patients’ journeys as there are no unified nationwide platforms that can consolidate all relevant data from all health care institutions in China [11,12]. The ever-increasing mobility of patients across the country has exacerbated the issue of data segmentation. Presumably, due to the recent improvement in annual income per household and the deployment of interstate health care systems, many patients opt for top-tier hospitals regardless of the travel distance from their home. It is quite common for one patient to receive treatment from different hospitals at various stages, while the hospital systems remain disconnected. The lack of longitudinal data from such fragmented health services may also contribute to subpar care and survival outcomes [13,14].

The distinct feature of telehealth, in which this study is interested, is its use as an adjunct to traditional physical visits and face-to-face consultations, particularly for posttreatment care and continuity of care from a physician’s perspective. Much of the research on telehealth usage has been focused on patients as the user population. Williams and Shang [15] examined telehealth use among a low-income, minority population in the United States and found the use of telehealth varies based on race, employment status, identified gender, education level, and...
age. Acoba et al [16] studied racial disparities in cancer patients during telehealth visits and confirmed that satisfaction with the visit is different between races. Turner et al [17] evaluated the experiences of health care providers and professionals during the COVID-19 pandemic and concluded the need for implementation strategies and necessary policies. Specifically for cancer patients, teledermatology has emerged as a popular mechanism [18]. In a cross-sectional study, Lama et al [19] found that more than one-half of cancer survivors use the internet or telehealth to access providers.

One of the specific aspects being assessed is follow-up, a unique challenge for cancer care in China, primarily due to the substantial patient-to-physician ratio [20]. Follow-ups using patient-reported outcomes (PROs) can improve the overall survival rate due to early relapse detection and better performance status at relapse. A study published in 2017 found that patients who reported their symptoms via an online tool survived 7 months longer than those who received usual care through regular screenings [21]. A previous meta-analysis of 21 studies also demonstrated that the reporting of PROs, including quality of life and disease symptoms, were significantly associated with tumor response to anticancer therapies such as chemotherapy, targeted therapy, and radiotherapy [22].

The platform used in this study, named WeDoc, is cloud-based and currently focuses on lung cancer. It consists of a mobile app for physicians, a WeChat mini-program for patients, and a cloud-based data and analytical component serving as the back end. The platform contains comprehensive, longitudinal medical records sourced from all relevant hospitals and supplemented with third-party test results, PROs, follow-up data, and more. The underlying data model is highly customizable to individual physicians’ needs and contains curated fields commonly used for cancer clinical research.

**Methods**

**Overview**

A cloud-based telehealth platform was built and launched for licensed oncologists and their patients. Patient medical records were collected and curated into a proprietary lung cancer data model. Physician and patient activities are recorded on the platform. A survey containing qualitative and quantitative questions was conducted 20 months after launch. Descriptive statistics and regression analysis were conducted on the survey data.

Analysis was conducted on 2 sets of data: activities recorded on the platform and results from a usage survey. Both sets of data were gathered from the perspective of physicians, as the goal in the first stage of this platform is to function as an assistant for physicians.

**Platform Implementation and Recording of User Activity**

The back end of the platform features a data processing pipeline: data and process management interfaces; and cloud repositories for raw, curated, and research data. Original data are deidentified, masking all personal details. These data are then abstracted and reviewed by trained personnel, and the abstracted data are consolidated, checked for quality, and committed to the real-world data repository.

Patients are invited to the platform by their oncologists and can form a user group with family members or friends. Oncologists can invite physicians and caregivers to create a treatment group, facilitating remote collaboration and simplifying hospital transfers. Patient reminders, assessments, and symptom feedback are gathered, and any potential adverse events are escalated to the primary oncologist.

The system’s data model incorporates the schema of electronic medical records, patient outcome reports, and periodic progression assessments by physicians, with a primary focus on lung cancer data. Data abstraction and data quality assurance involve both manual processes and regularly executed algorithms.

**Survey Design and Questionnaire**

The platform records the number of registered users and their activities. In March 2023, about 20 months after launch, an online usage survey was carried out using a WeChat survey mini-program. The program was pushed to all registered users as a study advertisement. The survey consisted of both qualitative and quantitative questions. Instead of individual physicians, each treatment group leader was asked to compile the group’s experience and provide responses. This approach was taken because the group leader dictates the use of the platform, and each group member may only utilize a subset of its functions.

The survey questions were designed to evaluate physicians’ patient care experiences using the platform. This includes basic functions and follow-up, their estimation of promptness in answering patient questions, patient risk reduction, cross-organization collaboration, and handling out-of-town patients. Quantitative questions asked for the number of both outpatients and inpatients managed, reduction in the number of physical visits, patient drop-off rates, and time spent collecting medical history during each visit. All identifiable information about participants was removed, and each individual was assigned a unique participant ID.
<table>
<thead>
<tr>
<th>Question category and description</th>
<th>Answer options</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient care functions</strong></td>
<td>Single-choice selection of binary options (agree or disagree) for each statement as a checkbox selection</td>
</tr>
<tr>
<td>A1: (Efficient communication) The platform serves as a communication channel for physicians to provide online notification of important matters.</td>
<td></td>
</tr>
<tr>
<td>A2: (Remote supervision) The platform enables physicians to provide continuous supervision and remote interaction.</td>
<td></td>
</tr>
<tr>
<td>A3: (Medical history retrieval) The platform offers patients’ medical history and communication records for physicians to review.</td>
<td></td>
</tr>
<tr>
<td>A4: (Patient administrative processes) The platform helps hospital appointment scheduling for both outpatient and inpatient procedures.</td>
<td></td>
</tr>
<tr>
<td>A5: (Response to patient question on time) The platform enables physicians to promptly answer patients’ questions without in-person visits.</td>
<td></td>
</tr>
<tr>
<td>A6: (Adverse event prevention) The platform enables physicians to timely capture potential adverse reactions from patient feedback.</td>
<td></td>
</tr>
<tr>
<td><strong>Follow-up</strong></td>
<td>Single-choice selection of binary options (agree or disagree) for each statement as a checkbox selection</td>
</tr>
<tr>
<td>B1: (Treatment status availability) Before: It was hard to acquire patient status. After: Patient status is easy to gather from the platform.</td>
<td></td>
</tr>
<tr>
<td>B2: (Survival status availability) Before: It was hard to acquire survival status. After: Survival status is provided on the platform.</td>
<td></td>
</tr>
<tr>
<td>B3: (Data comprehensiveness) Before: Records were incomplete. After: Multidimensional, comprehensive data are available on the platform.</td>
<td></td>
</tr>
<tr>
<td>B4: No differences between before and after using the platform.</td>
<td></td>
</tr>
<tr>
<td><strong>Response promptness</strong></td>
<td>Single-choice selection of 3 options (yes, no, or unknown) for each question as a radio button selection</td>
</tr>
<tr>
<td>C: With the platform, are you able to respond to patient inquiries quicker than before?</td>
<td></td>
</tr>
<tr>
<td><strong>Patient risk reduction</strong></td>
<td>Single-choice selection of 3 options (yes, no, or unknown) for each question as a radio button selection</td>
</tr>
<tr>
<td>D: After using the platform, do you feel that your patients have a lower risk of adverse reactions?</td>
<td></td>
</tr>
<tr>
<td><strong>Cross-organization collaboration</strong></td>
<td>Single-choice selection of 3 options (yes, no, or unknown) for each question as a radio button selection</td>
</tr>
<tr>
<td>E: Have you established collaborations across different departments, hospitals, or even regions through the platform?</td>
<td></td>
</tr>
<tr>
<td><strong>Management of remote patients</strong></td>
<td>Single-choice selection of 3 options (yes, no, or unknown) for each question as a radio button selection</td>
</tr>
<tr>
<td>F: Is managing out-of-town patients more convenient for you by using the platform?</td>
<td></td>
</tr>
<tr>
<td><strong>More patients managed per unit time</strong></td>
<td>Single-choice selection of 5 quantitative ranges: 10%-20%, 20%-50%, 50%-100%, &gt;100%, 0%</td>
</tr>
<tr>
<td>G: With the platform, how many more patients can you manage within the same amount of time?</td>
<td></td>
</tr>
<tr>
<td><strong>In-person visits saved</strong></td>
<td>Single-choice selection of 5 quantitative ranges: 1-3, 4-6, 7-10, &gt;11, 0</td>
</tr>
<tr>
<td>H: After using the platform, what is your estimation of the average number of in-person visits reduced per patient per year?</td>
<td></td>
</tr>
<tr>
<td><strong>Prior telehealth experience</strong></td>
<td>Single-choice selection of yes or no</td>
</tr>
<tr>
<td>I: Before using WeDoc, did you use any other telehealth platforms for patient management?</td>
<td></td>
</tr>
<tr>
<td><strong>Patient management specifics</strong></td>
<td>Quantitative values entered by participants</td>
</tr>
<tr>
<td>How many minutes per day do you spend managing patients?</td>
<td>Quantitative values entered by participants</td>
</tr>
<tr>
<td>J1: Before</td>
<td></td>
</tr>
<tr>
<td>J2: After</td>
<td></td>
</tr>
<tr>
<td><strong>What is the total number of patients you manage?</strong></td>
<td></td>
</tr>
</tbody>
</table>
Question category and description | Answer options
---|---
K1: Before | Quantitative values entered by participants
K2: After | Quantitative values entered by participants

**Outpatient management**

How many outpatient visits in total do your lung cancer patients have per month?
- L1: Before | Quantitative values entered by participants
- L2: After | Quantitative values entered by participants

What percentage of your lung cancer patients are likely to miss their outpatient visits each month?
- M1: Before | Quantitative values entered by participants
- M2: After | Quantitative values entered by participants

**Inpatient management**

How many lung cancer patients do you see for inpatient treatment per month?
- N1: Before | Quantitative values entered by participants
- N2: After | Quantitative values entered by participants

What percentage of your inpatients discontinue their treatment each month?
- O1: Before | Quantitative values entered by participants
- O2: After | Quantitative values entered by participants

**Medical history collection**

How many minutes do you spend collecting the medical history in each patient visit?
- P1: Before | Quantitative values entered by participants
- P2: After | Quantitative values entered by participants

**Statistical Analysis**

Descriptive statistics and regression analysis were conducted using the Python program. For descriptive analysis, we calculated the means, medians, standard deviations, and ranges. For quantitative questions regarding usage before and after, we used the Shapiro-Wilk test to assess the normal distribution of the data. Subsequently, we used the Wilcoxon rank sum test to evaluate the significance of the data sets. We used G*Power [23] to analyze the difference between 2 dependent means (matched pairs), setting the alpha at .05, beta at .2, and dz at 0.5. Assuming a medium-level difference between the before and after groups, a sample size of 27 was considered sufficient for the tests.

**Ethical Considerations**

This study was reviewed and approved by Yinchuan Ningfei Internet Hospital (approval number HLWYJ-2022-016). Participants were not compensated for their participation.

**Results**

**Activities Recorded on the Platform**

Over a period of 15 months, 608 physicians from 153 hospitals registered on the platform. The hospitals were from 21 of the 34 total provinces in China. Of the physicians, 92.8% (142/153) were from hospitals rated as Grade III, Level A, which is the highest rating according to the latest statistics [24] (Table 2). From a departmental perspective, 46.3% (125/270) of the physicians were from the oncology department, 41.9% (113/270) were from the department of respiratory and critical care medicine, and 11.9% (32/270) were from other departments.
Table 2. Physician and patient profiles registered in the system, including the numbers of hospitals, departments, physicians, treatment groups, and patients.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hospitals (n=153), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Grade III, Level A</td>
<td>142 (92.8)</td>
</tr>
<tr>
<td>Others</td>
<td>11 (7.2)</td>
</tr>
<tr>
<td><strong>Departments (n=270), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Oncology</td>
<td>125 (46.3)</td>
</tr>
<tr>
<td>Respiratory and critical care medicine</td>
<td>113 (41.9)</td>
</tr>
<tr>
<td>Others</td>
<td>32 (11.9)</td>
</tr>
<tr>
<td><strong>Physicians (n=608), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Treatment group leader</td>
<td>174 (28.6)</td>
</tr>
<tr>
<td>Treatment groups, n</td>
<td>211</td>
</tr>
<tr>
<td>Patients and family members, n</td>
<td>15,427</td>
</tr>
<tr>
<td><strong>Patients (n=13,888), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Nonresident patients</td>
<td>7826 (56.3)</td>
</tr>
</tbody>
</table>

One of the platform’s features for physicians is creating treatment groups by including other physicians. Among the 608 physicians, 174 have one or more groups. There are a total of 211 groups, with most physicians managing between 1 and 3 groups. A patient may be part of multiple groups, depending on the group’s purpose and treatment stage. For instance, a patient undergoing inpatient chemotherapy might initially be in a group with a radiologist in the hospital but later transferred to a follow-up group consisting only of the lead oncologist and the follow-up assistant. Table 2 describes the profiles of physicians and treatment groups.

Table 3. Activity log of the message types and quantities between physician-patient communication.

<table>
<thead>
<tr>
<th>Message type</th>
<th>Typical usage</th>
<th>Message count, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>Chats between patients and physicians</td>
<td>222,012</td>
</tr>
<tr>
<td>Reminder</td>
<td>Appointments and preparation items for appointments</td>
<td>66,985</td>
</tr>
<tr>
<td>Picture</td>
<td>Pictures in chat with patients</td>
<td>32,548</td>
</tr>
<tr>
<td>Patient education</td>
<td>General patient education through formats such as articles, videos, and URLs</td>
<td>27,538</td>
</tr>
<tr>
<td>Team message</td>
<td>Messages between physicians within the same group</td>
<td>19,779</td>
</tr>
<tr>
<td>Scaled assessment</td>
<td>Patient self-assessment of various aspects</td>
<td>8005</td>
</tr>
<tr>
<td>Voice</td>
<td>Voice messages for patients</td>
<td>5884</td>
</tr>
</tbody>
</table>

Survey Questionnaire Responses

Participant Characteristics

A total of 78 group leaders participated in the survey, representing 44.8% (78/174) of the treatment groups. All the groups were associated with Group III, Level A hospitals. Participant characteristics including city locations, gender distribution, departments, age groups, and prior experience with telehealth platforms are summarized in Table 4.
Table 4. Profiles of participants in the survey questionnaire (N=78).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Results, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>City location</strong></td>
<td></td>
</tr>
<tr>
<td>Beijing, Shanghai, or Guangzhou</td>
<td>33 (42)</td>
</tr>
<tr>
<td>Others</td>
<td>45 (58)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>36 (46)</td>
</tr>
<tr>
<td>Male</td>
<td>42 (54)</td>
</tr>
<tr>
<td><strong>Departments</strong></td>
<td></td>
</tr>
<tr>
<td>Oncology</td>
<td>46 (59)</td>
</tr>
<tr>
<td>Respiratory</td>
<td>26 (33)</td>
</tr>
<tr>
<td>Others</td>
<td>6 (8)</td>
</tr>
<tr>
<td><strong>Age group (years)</strong></td>
<td></td>
</tr>
<tr>
<td>20-30</td>
<td>4 (5)</td>
</tr>
<tr>
<td>30-40</td>
<td>15 (19)</td>
</tr>
<tr>
<td>40-50</td>
<td>33 (42)</td>
</tr>
<tr>
<td>50-60</td>
<td>24 (31)</td>
</tr>
<tr>
<td>&gt;60</td>
<td>2 (3)</td>
</tr>
<tr>
<td><strong>Prior telehealth usage</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>25 (32)</td>
</tr>
<tr>
<td>Yes</td>
<td>53 (68)</td>
</tr>
</tbody>
</table>

**Qualitative Question Results**

For questions A1 to F, which included the topics of communication efficiency, remote supervision, question response times, adverse event prevention, follow-up data completeness, patient risk reduction, cross-organization collaboration, and remote patient management, participants provided qualitative answers to each question. The results are shown in Table 5. A positive answer indicates agreement with the statement or yes to the question. A negative answer indicates disagreement with the statement or no to the question. Most of the questions received positive answers except for the topic of cross-organization collaboration, which had nearly neutral feedback: 54% positive versus 46% negative. The questions of treatment status availability (B1), survival status availability (B2), and data comprehensiveness (B3) contain both before and after statements. A negative answer may indicate that the participant only disagrees with part of the statement. Therefore, the final results of these questions indicated less favorable evaluations of the WeDoc tool.
<table>
<thead>
<tr>
<th>Question description</th>
<th>Survey results, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A1: Efficient communication</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>76 (97)</td>
</tr>
<tr>
<td>Negative</td>
<td>2 (3)</td>
</tr>
<tr>
<td><strong>A2: Remote supervision</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>73 (94)</td>
</tr>
<tr>
<td>Negative</td>
<td>5 (6)</td>
</tr>
<tr>
<td><strong>A3: Medical history retrieval</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>69 (89)</td>
</tr>
<tr>
<td>Negative</td>
<td>9 (12)</td>
</tr>
<tr>
<td><strong>A4: Patient administrative processes</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>58 (74)</td>
</tr>
<tr>
<td>Negative</td>
<td>20 (26)</td>
</tr>
<tr>
<td><strong>A5: Respond to patient questions on time</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>69 (89)</td>
</tr>
<tr>
<td>Negative</td>
<td>9 (12)</td>
</tr>
<tr>
<td><strong>A6: Adverse event prevention</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>67 (86)</td>
</tr>
<tr>
<td>Negative</td>
<td>11 (14)</td>
</tr>
<tr>
<td><strong>B1: Treatment status availability</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>68 (87)</td>
</tr>
<tr>
<td>Negative</td>
<td>10 (13)</td>
</tr>
<tr>
<td><strong>B2: Survival status availability</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>58 (74)</td>
</tr>
<tr>
<td>Negative</td>
<td>20 (26)</td>
</tr>
<tr>
<td><strong>B3: Data comprehensiveness</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>68 (87)</td>
</tr>
<tr>
<td>Negative</td>
<td>10 (13)</td>
</tr>
<tr>
<td><strong>B4: No difference</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>8 (10)</td>
</tr>
<tr>
<td>Negative</td>
<td>70 (90)</td>
</tr>
<tr>
<td><strong>C: Response promptness</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>70 (90)</td>
</tr>
<tr>
<td>Negative</td>
<td>8 (10)</td>
</tr>
<tr>
<td><strong>D: Patient risk reduction</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>70 (90)</td>
</tr>
<tr>
<td>Negative</td>
<td>8 (10)</td>
</tr>
<tr>
<td><strong>E: Cross-organization collaboration</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>42 (54)</td>
</tr>
<tr>
<td>Negative</td>
<td>36 (46)</td>
</tr>
<tr>
<td><strong>F: Management of remote patients</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>78 (100)</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
</tr>
</tbody>
</table>
Quantitative Question Results

For questions J to P, participants were asked to provide quantitative values for their experiences both before and after using the tool. Table 6 summarizes the values for each question. The “Unknown” category indicates null values in the survey, and these responses were omitted in the analysis. The highest number of unknown answers we received was for the question about the number of patients managed before using the tool.

We used G*Power analysis for the remaining nonnull before-and-after pairs to ensure that there was a sufficient sample for analysis. With an assumption of medium differences between the before and after groups, at least 27 samples had to be present in the group.
### Table 6. Results for the quantitative survey questions (N=78).

<table>
<thead>
<tr>
<th>Question description</th>
<th>Unknown responses, n (%)</th>
<th>Valid responses</th>
<th>Minimum-maximum</th>
<th>Median</th>
<th>Mean (SD)</th>
<th>Mean improvement, %</th>
<th>P value for the before-after comparison&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time spent managing patients (minutes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J1: Before</td>
<td>18 (23)</td>
<td>10-180</td>
<td>30</td>
<td>50.5 (45.1)</td>
<td>N/A&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>J2: After</td>
<td>10 (13)</td>
<td>3-120</td>
<td>20</td>
<td>25.5 (22.7)</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of patients managed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.01</td>
</tr>
<tr>
<td>K1: Before</td>
<td>20 (26)</td>
<td>0-800</td>
<td>40</td>
<td>105.7 (177.4)</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K2: After</td>
<td>13 (17)</td>
<td>1-1606</td>
<td>100</td>
<td>324.3 (428.8)</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Monthly number of outpatient lung cancer patients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.66</td>
</tr>
<tr>
<td>L1: Before</td>
<td>8 (10)</td>
<td>0-2000</td>
<td>85</td>
<td>221.8 (352.2)</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2: After</td>
<td>8 (10)</td>
<td>0-2000</td>
<td>80</td>
<td>237.1 (369.6)</td>
<td>___c</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Outpatient drop-off rate (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>M1: Before</td>
<td>11 (14)</td>
<td>0-80</td>
<td>25</td>
<td>26.8 (21.1)</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2: After</td>
<td>11 (14)</td>
<td>0-50</td>
<td>10</td>
<td>13.1 (11.4)</td>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Monthly number of lung cancer inpatients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.59</td>
</tr>
<tr>
<td>N1: Before</td>
<td>2 (3)</td>
<td>0-350</td>
<td>60</td>
<td>110.1 (93.5)</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2: After</td>
<td>2 (3)</td>
<td>0-350</td>
<td>70</td>
<td>116.2 (94.8)</td>
<td>___</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Inpatient drop-off rate (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>O1: Before</td>
<td>5 (6)</td>
<td>0-50</td>
<td>10</td>
<td>14.9 (12.1)</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2: After</td>
<td>5 (6)</td>
<td>0-100</td>
<td>5</td>
<td>9.2 (16.2)</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time collecting medical history (minutes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>P1: Before</td>
<td>5 (6)</td>
<td>2-180</td>
<td>10</td>
<td>14.0 (28.8)</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2: After</td>
<td>3 (4)</td>
<td>1-120</td>
<td>3</td>
<td>8.8 (19.2)</td>
<td>57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Assessed using Wilcoxon tests.

<sup>b</sup>N/A: not applicable.

<sup>c</sup>No improvement.

To better understand the differences between the before and after results, we used the Shapiro algorithm to test whether the values fell within a normal distribution. For normally distributed data series, a t test can be used to compare the pairs. Otherwise, the Wilcoxon test is a more suitable method. Since the P values of the Shapiro test were all <.001, which is much lower than the common hypothesis threshold of .05, we concluded that none of the pairs were normally distributed. Therefore, Wilcoxon tests were performed on the before-and-after pair data (Table 6). The Wilcoxon results suggest that there were 2 questions that were not significantly different between before and after the platform: the monthly number of outpatients admitted and the monthly number of inpatients admitted. This result is quite explainable, as the telehealth tool itself is not aimed at recruiting new patients; therefore, the monthly numbers of patients remained nearly the same. For the topics that had significant changes, we calculated the improvements based on the mean values collected in the survey, which are also shown in Table 6.

Although the survey was not specifically designed to compare the group with prior telehealth platform experience with the group without prior experience, we discovered that 68% (53/78) of the participants had used telehealth tools before. In order to understand the experience by group, we carried out a Wilcoxon test to compare the responses before and after (Table 7). The numbers of monthly admitted outpatient and inpatient lung cancer patients still did not change significantly. However, there was also no significant change in the number of patients managed, suggesting that physicians may not manage more patients using WeDoc than with other telehealth platforms.
Table 7. Results for the quantitative survey questions for those who had prior telehealth platform experience (n=53).

<table>
<thead>
<tr>
<th>Question description</th>
<th>Unknown responses, n (%)</th>
<th>Minimum-maximum</th>
<th>Median</th>
<th>Mean (SD)</th>
<th>Mean improvement, %</th>
<th>P value for the before-after comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time spent managing patients (minutes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J1: Before</td>
<td>5 (9)</td>
<td>10-180</td>
<td>30</td>
<td>50.7 (46.8)</td>
<td></td>
<td>N/A b</td>
</tr>
<tr>
<td>J2: After</td>
<td>5 (9)</td>
<td>3-120</td>
<td>20</td>
<td>28.6 (25.2)</td>
<td>77</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Number of patients managed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.06</td>
</tr>
<tr>
<td>K1: Before</td>
<td>11 (21)</td>
<td>1-800</td>
<td>40</td>
<td>125.8 (199.5)</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>K2: After</td>
<td>6 (11)</td>
<td>1-1200</td>
<td>100</td>
<td>322.3 (403.4)</td>
<td></td>
<td>_ c</td>
</tr>
<tr>
<td>Monthly number of outpatient lung cancer patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.75</td>
</tr>
<tr>
<td>L1: Before</td>
<td>4 (8)</td>
<td>0-2000</td>
<td>60</td>
<td>214.7 (388.9)</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>L2: After</td>
<td>4 (8)</td>
<td>0-2000</td>
<td>60</td>
<td>225.6 (407.9)</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Outpatient drop-off rate (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>M1: Before</td>
<td>7 (13)</td>
<td>0-80</td>
<td>30</td>
<td>28.2 (22.7)</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>M2: After</td>
<td>7 (13)</td>
<td>0-50</td>
<td>10</td>
<td>13.9 (12.1)</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>Monthly number of lung cancer inpatients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.65</td>
</tr>
<tr>
<td>N1: Before</td>
<td>0</td>
<td>0-350</td>
<td>60</td>
<td>109.6 (100.0)</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>N2: After</td>
<td>0</td>
<td>0-350</td>
<td>60</td>
<td>113.8 (98.4)</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Inpatient drop-off rate (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>O1: Before</td>
<td>3 (6)</td>
<td>0-50</td>
<td>10</td>
<td>14.7 (13.3)</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>O2: After</td>
<td>3 (6)</td>
<td>0-12</td>
<td>5</td>
<td>5.7 (4.1)</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>Time collecting medical history (minutes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>P1: Before</td>
<td>3 (6)</td>
<td>2-180</td>
<td>8</td>
<td>15.3 (34.1)</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>P2: After</td>
<td>3 (6)</td>
<td>1-120</td>
<td>3</td>
<td>9.3 (22.8)</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

a Assessed using Wilcoxon tests.
b N/A: not applicable.
c No improvement.

Discussion

Principal Findings

Specific Feedback About the Platform

Results from activity logs and survey responses demonstrate the feasibility of cancer patient care using telehealth with a live-transmitted real-world database. Specifically, 84% (65.6/78, SD 8.7) of participants responded positively to questions A1 through F. The lowest scores were for patient administrative processes and survival status. Patient administrative processes in China are complex and not the primary focus of this platform, while obtaining updated survival status during follow-up is clearly an area for improvement. Another area that did not stand out was cross-organization collaboration, presumably due to the deployment of other specialized platforms such as Multidisciplinary Team, which is popular in China. Of the participants, 92% believed that they could manage more patients with the same amount of time, and an equal number of physicians agreed that the platform saves at least one or more instances of in-person visits.

Our analysis of the before and after experiences of the same population showed that 5 of the 7 categories were significantly different after use of the platform, as determined using the Wilcoxon signed rank test. The 2 categories that were not significantly different were the monthly numbers of outpatient and inpatient admissions. These 2 factors are unaffected by the use of any patient management tool; thus, they are indeed irrelevant to our telehealth platform.

Perceptions of Those With Prior Telehealth Usage

Given that 68% of the participants had prior experience with telehealth platforms, analyzing this population alone yielded similar results, except that the number of patients managed did not meet our significance value assumption of .05. This implies that, although managing more patients is a benefit of telehealth platforms, it may not be unique to ours. The strengths of a telehealth platform with real-world data are manifested in the
categories of time efficiency, drop-off rates, and access to patients’ medical histories.

**Remote Patient Management**

The adoption of remote patient management was evident in the patient profiles, which showed that more than one-half of patients, about 56.3%, were nonresidents; 941 patients had transferred from one hospital to another, and almost 1500 patients had prior diagnoses or treatments from hospitals other than their current hospital. Taking hospitals in Shanghai as an example, the platform showed that about 35% of patients were from cities other than Shanghai. Although more than one-half of the patients were from adjacent provinces such as Jiangsu and Zhejiang, some travel thousands of miles from places like Heilongjiang, Sichuan, and Liaoning. Because of the unbalanced health care situation in China, it is quite common for patients to be diagnosed in one hospital and receive treatments at another. Despite significant improvements over the past few decades, the best oncologists and medical facilities are still heavily concentrated in top cities.

**Text as the Dominant Message Type**

The activity log indicated that text was the most commonly used message type to communicate with patients. The use of pictures and voices messages was significantly lower than that of text. Reminders were also quite popular, followed by educational materials. The preliminary analysis did not reveal significant differences in usage patterns among physicians, so we did not present usage data by physician profile.

**Security and Privacy**

With the adoption of the Personal Information Protection Law (PIPL) [25] in China on November 01, 2021, all systems handling data from Chinese citizens must be compliant with the law. This law is widely seen as China’s equivalent of the EU General Data Protection Regulation (GDPR) [26]. The system in question acts as both a data handler and data processor. It controls the scope of data usage based on the level of consent obtained from users, making user consent a mandatory prerequisite for successful user registration. By separating raw data and identifiers from curated, deidentified data, the system ensures the proper implementation of data protection policies.

From an operational perspective, privacy protection remains one of the most significant challenges in building such a platform. The challenge is less technical, as there are rich sets of mechanisms available, such as data anonymization, encryption, access control, and audit. The main challenge comes from the perceptions and cooperation of patients. Ideally, patients and their relatives should also have access to real-world data, enabling them to participate in treatment decisions. Apart from patient perceptions, potential malpractice concerns also hinder data sharing, preventing people from gaining strategic insights. Health care policymakers and scientific researchers need to collaborate with data analysts to promote a proper data sharing process.

**Limitations**

Although this study is based on a live system with real-world data and experiences, the findings remain preliminary. At present, the platform only provides services to the lung cancer population, and the results of this study are derived from physicians from a subset of the treatment paradigm. Although the user base of the platform encompasses both physicians and patients, future research involving a broader population, including more physicians and direct patient experiences, may yield new, insightful findings. It would also be interesting to expand to other diseases. Given the large quantity of chat messages accumulated on the platform, a detailed examination of these messages paired with language processing models would be a fascinating next step.

**Conclusion**

This study demonstrates the feasibility of using telehealth for patient management. As the focus of cancer treatment shifts toward patient care, telehealth in the form of mobile apps, web-based interfaces, or other formats will play an increasingly critical role in enabling physicians to maintain close contact with patients, regardless of physical location. We advocate for the integration of telehealth with comprehensive real-world medical record data, so that such a platform can provide patient management capabilities. This could eventually lead to improved quality of life and survival rates of cancer patients.

**Data Availability**

The data sets analyzed during this study are available from the corresponding author on reasonable request.

**Authors’ Contributions**

DZ, YS, YR, and LJ contributed to the conception and design of the study. JN, LP, XT, ZD, YZ, AG, JW, XL, and JZ refined the research questions and provided feedback on the study design. WH, CZ, CL, HL, YD, JX, DW, XC, RM, and XD assisted with platform function review, usability testing, and participant recruitment. YS and SL designed the survey questions and analyzed the data. DZ, LP, XT, YD, XD, YR, and LJ drafted the manuscript.

**Conflicts of Interest**

YR and SL are co-founders of Metafame Technologies Inc, which developed the system. However, the system is not marketed as a paid service.
References


Abbreviations

GDPR: General Data Protection Regulation  
PIPL: Personal Information Protection Law  
PRO: patient-reported outcome 

©Di Zheng, Yanhong Shang, Jian Ni, Ling Peng, Xiaoming Tan, Zhaoxia Dai, Yizhuo Zhao, Aiqin Gu, Jiying Wang, Yanyan Song, Xiaofeng Li, Junping Zhang, Wei Heng, Cuiying Zhang, Chunling Liu, Hui Li, Yingying Du, Jianfang Xu, Dan Wu, Xuwei Cai, Rui Meng, Xiaorong Dong, Yaoping Ruan, Liyan Jiang. Originally published in JMIR Cancer (https://cancer.jmir.org), 05.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Applying the Unified Theory of Acceptance and Use of Technology to Identify Factors Associated With Intention to Use Teledelivered Supportive Care Among Recently Diagnosed Breast Cancer Survivors During COVID-19 in Hong Kong: Cross-Sectional Survey

Nelson C Y Yeung¹, BSocSci, MPhil, PhD; Stephanie T Y Lau¹, BSocSci; Winnie W S Mak², PhD; Cecilia Cheng³, BSocSci, PhD; Emily Y Y Chan¹, BS, MBBS, SM PH, MD; Judy Y M Siu⁴, BSocSci, MPhil, PhD; Polly S Y Cheung⁵, MBBS

¹JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
²Department of Psychology, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
³Department of Psychology, The University of Hong Kong, Hong Kong, China (Hong Kong)
⁴Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)
⁵Hong Kong Breast Cancer Foundation, Hong Kong, China (Hong Kong)

Corresponding Author:
Nelson C Y Yeung, BSocSci, MPhil, PhD
JC School of Public Health and Primary Care
The Chinese University of Hong Kong
Room 508, Postgraduate Education Centre
Prince of Wales Hospital
Hong Kong
China (Hong Kong)
Phone: 852 22528740
Email: nelsonyeung@cuhk.edu.hk

Abstract

Background: Many supportive cancer care (SCC) services were teledelivered during COVID-19, but what facilitates patients’ intentions to use teledelivered SCC is unknown.

Objective: The study aimed to use the unified theory of acceptance and use of technology to investigate the factors associated with the intentions of breast cancer survivors (BCS) in Hong Kong to use various types of teledelivered SCC (including psychosocial care, medical consultation, complementary care, peer support groups). Favorable telehealth-related perceptions (higher performance expectancy, lower effort expectancy, more facilitating conditions, positive social influences), less technological anxiety, and greater fear of COVID-19 were hypothesized to be associated with higher intentions to use teledelivered SCC. Moreover, the associations between telehealth-related perceptions and intentions to use teledelivered SCC were hypothesized to be moderated by education level, such that associations between telehealth-related perceptions and intentions to use teledelivered SCC would be stronger among those with a higher education level.

Methods: A sample of 209 (209/287, 72.8% completion rate) women diagnosed with breast cancer since the start of the COVID-19 outbreak in Hong Kong (ie, January 2020) were recruited from the Hong Kong Breast Cancer Registry to complete a cross-sectional survey between June 2022 and December 2022. Participants’ intentions to use various types of teledelivered SCC (dependent variables), telehealth-related perceptions (independent variables), and sociodemographic variables (eg, education, as a moderator variable) were measured using self-reported, validated measures.

Results: Hierarchical regression analysis results showed that greater confidence using telehealth, performance expectancy (believing telehealth helps with daily tasks), social influence (important others encouraging telehealth use), and facilitating conditions (having resources for telehealth use) were associated with higher intentions to use teledelivered SCC (range: β=0.16, P=.03 to β=0.34, P<.001). Moreover, 2-way interactions emerged between education level and 2 of the telehealth perception variables. Education level moderated the associations between (1) performance expectancy and intention to use teledelivered complementary care (β=0.34, P=.04) and (2) facilitating conditions and intention to use teledelivered peer support groups (β=0.36,
Conclusions: The findings of this study implied that enhancing BCS’ skills at using telehealth, BCS’ and their important others’ perceived benefits of telehealth, and providing assistance for telehealth use could increase BCS’ intentions to use teledelivered SCC. For intentions to use specific types of SCC, addressing relevant factors (performance expectancy, facilitating conditions) might be particularly beneficial for those with a higher education level.

Key words
telehealth; tele-delivered supportive cancer care; breast cancer; COVID-19; technology acceptance; UTAUT

Introduction

Potential Impacts of COVID-19 on Breast Cancer Care

The COVID-19 pandemic has been an international public health emergency, posing severe threats to lives and health care systems worldwide. In Hong Kong, the implementation of different preventive measures (eg, regulations for social distancing, reprioritization of hospital services) affected the lives of not only the general population but also individuals with chronic diseases. Being one of the most commonly diagnosed cancers in Hong Kong, breast cancer diagnosis and treatment delays occurred during the COVID-19 pandemic [1]. For example, the number of pathologic specimens for the 4 most common cancer regions in Hong Kong (including breast cancer) received by public laboratories and public hospitals for cancer diagnostic services reduced by 15.5% overall in 2020, compared with the prior 3-year average [2]. Another study suggested that breast cancer patients in Hong Kong needed to wait 3 weeks longer for their first specialist consultation during the COVID-19 crisis than before the pandemic [3].

After completion of active treatments, many breast cancer survivors (BCS) still need supportive cancer care (SCC) and rehabilitation services to help with different cancer-related life aspects [4]. In the Netherlands, one-third of 1051 surveyed BCS reported difficulties contacting their general practitioner due to COVID-19 [5]. The COVID-19-related lockdowns in the United States and Germany also disrupted patients’ referrals to cancer survivorship programs [6,7]. To reduce the impact of COVID-19 on cancer care, alternative modes of SCC delivery are therefore important.

Acceptability of Telehealth for Cancer Patients

Research suggests that COVID-19 might have catalyzed new models of health care (eg, telehealth) [4]. Telehealth is the use of technology to deliver health care, health information, or health education at a distance [8]. Telehealth technologies (including telephone, videoconferencing, and internet-based intervention) can bring services into the patient’s home and help them cope with their illness without the need to be physically present at a hospital or clinic [8]. A recent qualitative study in Australia reported that patients with hematological cancer considered telehealth an acceptable alternative during the pandemic [9]. However, some patients encountered difficulties using teledelivered cancer care services due to a lack of knowledge and skills, plus some preferred to see the doctor visually through a video call over other teledelivered options [9]. Another survey explored the prospect of using telemedicine for follow-up among Australian BCS and found that 70% of respondents had suitable devices to access telehealth but only 15% accepted the postoperation teleconsultation with their surgeon [10]. Given that relevant research is limited in the Hong Kong context, this study examined the level of acceptability of telehealth for BCS to access SCC and its associated factors amid the COVID-19 pandemic.

Telehealth-Related Perceptions as Determinants of Patients’ Intentions to Use Telehealth for SCC

Different theoretical models have been applied to explore intentions to use telehealth among general healthy populations and patient populations outside the COVID-19 context [11]. Among the models, the unified theory of acceptance and use of technology (UTAUT) is one of the most influential theories to understand people’s acceptance of different types of information technologies including telehealth [11]. According to the UTAUT, performance expectancy (whether the individuals believe using the system would provide benefits), effort expectancy (whether the system is easy to use), social influence (perception of important others’ opinions about using the system), facilitating conditions (organizational and technical infrastructure supporting the use of the system), and technology anxiety (users’ negative emotional states related to learning to use technology [eg, nervousness, fear]) are the important determinants of people’s intentions to use technology [12]. Compared with other traditional behavioral theories (eg, Theory of Planned Behavior, Health Belief Model), the UTAUT seems to have stronger explanatory power for understanding people’s intentions to use telehealth [11].

The model has been applied to people’s use of telehealth in different disease contexts. For example, higher performance expectancy, lower effort expectancy, more favorable social influences, less technology anxiety, and more facilitating factors have been associated with intention to use telehealth among Chinese populations (eg, older individuals in the community, individuals with chronic diseases) [13,14]. Performance expectancy and social influence were associated with higher intention to use telehealth service and treatment among patients with diabetes in Korea [15]. Similarly, among patients with type 2 diabetes in South Africa [16], lower performance expectancy, lower effort expectancy, less social influence, and fewer facilitating conditions explained the generally lower intention to use telehealth services. To the best of our knowledge, research on examining cancer survivors’ intentions to use teledelivered
SCC during the COVID-19 pandemic was limited. Therefore, this study aimed to examine how telehealth-related perceptions were associated with intentions to use telehealth for SCC among BCS in Hong Kong during the COVID-19 pandemic.

### Individual Characteristics and Fear of COVID-19 as Potential Determinants of Intentions to Use Telehealth for Supportive Cancer Services Among BCS

In addition to telehealth-related perceptions, patients’ sociodemographic characteristics might also contribute to the acceptability of telehealth [11]. Factors like age, education, possession of smart device(s), the nature of the consultation (routine follow-up versus urgent need for physical examination), and experience with using technology could contribute to the acceptability of telehealth for cancer survivors [17]. Specific to the pandemic situation, recent studies found that fear of COVID-19 transmission was associated with higher intentions to use contact tracing apps among the general population in Germany [18] and telehealth services among cancer patients in the United States [19]. Expecting the same phenomenon to apply to BCS in Hong Kong, we aimed to examine the roles of patients’ individual characteristics (eg, sociodemographic and clinical factors, fear of COVID-19) and prior experience with using technology in intentions to use telehealth for SCC.

### Moderating Role of Education Level

Despite the wide use of the UTUAT to explain people’s intentions to use technology, whether the contribution of the variables in the theory differs based on people’s sociodemographic and individual characteristics has not been extensively examined. Prior studies have generally regarded sociodemographic variables as covariates for intentions or behavior, which fails to unpack the complex ways in which such characteristics might interact with beliefs to determine behavioral intention and actual behaviors (eg, [20-22]). Education level has been suggested as a potential moderator between perceptions about behaviors and intentions to engage in online behaviors. For example, studies measured the intention of individuals to use e-banking based on the UTAUT model in the United Kingdom and Jordon and found that education level had a positive moderating effect on performance expectancy, facilitating conditions [23], and effort expectancy [24]. Another study in Indonesia also found that education level moderated the relationship between effort expectancy and intention to use e-money services [25]. Similar research on the intentions of BCS to use telehealth amid the COVID-19 pandemic was limited. Specifically, the role of education as a moderator between telehealth perceptions and BCS’ intentions to use teledelivered SCC were investigated in this study.

### Purpose of the Study

This study aimed to examine how telehealth-related perceptions contribute to the intention to use telehealth for cancer care among BCS in Hong Kong during the COVID-19 pandemic (Figure 1). We hypothesized that favorable telehealth-related perceptions (higher performance expectancy, lower effort expectancy, more facilitating conditions, positive social influences), less technological anxiety, and greater fear of COVID-19 would be associated with higher intention to use telehealth for SCC. We also hypothesized that the associations between telehealth-related perceptions and intentions to use teledelivered SCC would be moderated by education level, such that associations between telehealth-related perceptions (higher performance expectancy, lower effort expectancy, more facilitating conditions, positive social influences, less technological anxiety) and intention to use teledelivered SCC would be stronger among those with a higher education level.

### Methods

#### Participants and Procedure

A cross-sectional study was conducted. BCS were eligible to participate if they (1) were older than 18 years, (2) had a confirmed diagnosis of Stage 0-III breast cancer since the outbreak of COVID-19 in Hong Kong (January 2020), (3) were receiving active treatment (eg, radiotherapy, chemotherapy), (4) could read Chinese to answer questionnaires and communicate in Cantonese, and (5) were able to provide meaningful informed consent. BCS were excluded if they had (1) a history of any psychiatric disorder, (2) metastatic brain
disease, (3) any other type of cancer, or (4) recurrent breast cancer.

Prospective participants were recruited from the Hong Kong Breast Cancer Registry (HKBCR). The HKBCR has been the most comprehensive, representative local data collection and monitoring system for BCS in Hong Kong [26]. Upon approval, BCS who fulfilled the inclusion criteria based on the data in the HKBCR were invited to participate in the study through telephone calls. Of the 943 BCS contacted, 409 were not reachable, 23 were not eligible, and 227 were not interested in the study. With initial verbal consent via phone, those who were eligible and interested in the study (N=287) were asked to complete the cross-sectional survey. Participants received a cover letter explaining the study details, consent form, packet of questionnaires, stamped return envelope, thank you/reminder letter, and replacement packet via mail. After consent, participants completed the survey in the home setting. Telephone calls were used to remind individuals who had not returned the questionnaires. The study was conducted between June 2022 and December 2022 (amid the fifth wave of the COVID-19 pandemic in Hong Kong) [27]. A total of 209 completed surveys were returned (out of 287 sent), yielding a completion rate of 72.8%.

Ethical Considerations
Ethics approvals were sought from the Joint Chinese University of Hong Kong - New Territories East Cluster Clinical Research Ethics Committee (CREC Ref. 2021.286) and Hong Kong Breast Cancer Foundation. We obtained informed consent before participation in the survey. Upon completion of the survey, participants received supermarket vouchers (worth HK$100; approximately US $12.80) to compensate them for their time. We guaranteed that the identity of the participants would not be revealed.

Sample Size Planning
The dependent variable was the intention to use teledelivered SCC services. Based on prior studies on the acceptability of telehealth among Chinese populations [28,29], we expected a small to medium overall effect size ($\alpha^2 =0.10$) in the association between telehealth-related perceptions and intentions to use telehealth services in the hierarchical regression analysis. To achieve a statistical power of .80 at $\alpha =0.05$, a minimum of 201 participants were needed (G*Power 3.1.2). The sample size (N=209) achieved via the recruitment strategy was expected to allow the detection of the expected effect size with sufficient statistical power.

Measures
A written, closed-ended, anonymous, self-administered questionnaire was used in the study. To ensure that the questionnaire was readily comprehensible, a pilot test was conducted among 10 BCS who were eligible for the study. The study questionnaire was finalized based on feedback from the pilot test participants.

Intention to Use Telehealth for Future Supportive Cancer Services
Participants’ intentions to use telehealth for future supportive cancer services was measured using a SCC service utilization scale [30] that was modified according to the local health care context. The checklist covered different categories of services, including psychological support (6 items; $\alpha =.91$), medical consultation (5 items; $\alpha =.86$), integrated or complementary care (6 items; $\alpha =.87$), and peer support (2 items; $\alpha =.83$). On a 4-point scale (1, no intention or not applicable; 2, low intention; 3, moderate intention; 4, high intention), participants were asked to indicate their intention to use telehealth for each SCC service (e.g., “I intend to use telehealth for psycho-oncology counseling.”). The scale has been shown to be reliable and valid among Western cancer survivors [30].

Perceptions About Telehealth for SCC Services
We used 4 subscales (performance expectancy [3 items], effort expectancy [4 items], social influence [3 items], and facilitating conditions [3 items]) to measure participants’ perceived usefulness, perceived ease, social influence, and facilitating conditions, respectively, for using telehealth in cancer care [31]. Sample items include “Using telehealth for cancer care is beneficial to my health.” ($\alpha =.83$; performance expectancy), “It is easy for me to become skillful at using telehealth for cancer care service.” ($\alpha =.87$; effort expectancy), “People whose opinions that I value (eg, my doctors) think I should use telehealth for cancer care services.” ($\alpha =.86$; social influence), and “I have the resources necessary to use telehealth for cancer care services.” ($\alpha =.90$; facilitating conditions). On a 5-point scale (1, strongly disagree; 5, strongly agree), higher mean item scores from the scales indicate higher levels of the corresponding constructs. The Chinese versions of these scales were shown to be reliable and valid among Chinese adults [32].

Technology Anxiety
A 3-item scale was adapted to measure participants’ technology anxiety while using telehealth services [14]. On a 5-point scale (1, strongly disagree; 5, strongly agree), a higher mean item score indicates a higher level of technology anxiety (e.g., “I feel nervous about using telehealth.” $\alpha =.91$). The Chinese version of the scale was shown to be reliable and valid among Chinese adults [14].

Fear of COVID-19
The Chinese version of the 7-item Fear of COVID-19 scale was adapted to measure participants’ fear of COVID-19 [33]. On a 5-point scale (1, strongly disagree; 5, strongly agree), a higher mean item score indicates a higher level of COVID-19 fear (e.g., “It makes me uncomfortable to think about COVID-19.” $\alpha =.88$). The scale has been shown to be reliable and valid in the Chinese population [34].

Clinical and Sociodemographic Characteristics
Participants self-reported their (1) sociodemographic characteristics (e.g, age, education level, employment status, marital status), (2) treatment-related variables (e.g, surgeries undergone, treatments receiving or undergone, time since last treatment), (3) daily living variables (e.g, access to the internet,
Cancer Care Experiences During COVID-19

Participants were asked if they had participated in any telehealth online consultation sessions for SCC (including psychological support services, medical support services, integrated and complementary support services, spiritual support services, other support services; no=0, yes=1).

Statistical Analysis

Descriptive and bivariate Pearson correlation analyses were conducted. Hierarchical regression analyses were also conducted to examine factors associated with intentions to use telehealth for supportive cancer services. The sequence of entering independent variables followed suggestions from prior studies that examined factors associated with people’s health or health behavior outcomes and the interaction effects among those factors (eg, [35,36]). The process usually involves entering important sociodemographic and individual experience variables in the first block (as a statistical control for confounding variables), variables representing major theoretical constructs in the next block(s), and the interaction terms between the proposed moderating variable and the independent variables of interest in the last block. In our study, fear of COVID-19 and the sociodemographic and clinical variables that had significant bivariate correlations with the dependent variables were entered in block 1 of the regression model. Telehealth-related perceptions (ie, performance expectancy, effort expectancy, social influence, facilitating conditions, technology anxiety) were entered into block 2 of the regression model. In the last block, 5 interaction terms between telehealth-related perceptions and education level were entered into the model. To compute the interaction terms, the mean-centered scores of telehealth perceptions and education level (binary: college level versus below college level) were multiplied. All continuous independent variables were centered prior to the analyses. For statistically significant interactions, simple slopes analyses [37] were conducted to examine how the main effects of telehealth perceptions on intentions to use teledelivered SCC varied at different education levels. Those with $P \leq .05$ in the final regression model were considered statistically significant. These analyses were performed using SPSS version 26.0.

Results

Participant Characteristics

Among the 209 participants, 82 (39.2%) were 50 years or younger, 63 (30.1%) were 51 years to 60 years old, and 62 (29.7%) were at least 61 years old. In addition, of the 209 participants, 91 (43.5%) had a tertiary education, 72 (34.4%) worked full-time, 99 (47.4%) reported a religious affiliation, and 53 (25.4%) had a comorbid chronic illness. Regarding cancer-related characteristics, 10 (4.8%), 60 (28.7%), 86 (41.1%), and 53 (25.4%) of the 209 participants reported being diagnosed with Stage 0, Stage I, Stage II, and Stage III breast cancer, respectively, and 194 (94.3%) had undergone breast cancer surgery. The average time since diagnosis was 16.6 (SD 8.00) months. Regarding internet access, 204 of the 209 participants (97.6%) had a mobile phone with internet access (Table 1).
Table 1. Demographic characteristics of the participants (N=209).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), n (%)</td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>82 (39.2)</td>
</tr>
<tr>
<td>51-60</td>
<td>63 (30.1)</td>
</tr>
<tr>
<td>≥61</td>
<td>62 (29.7)</td>
</tr>
<tr>
<td>Refused to answer</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Gender (female), n (%)</td>
<td>209 (100)</td>
</tr>
<tr>
<td>Cancer stage, n (%)</td>
<td></td>
</tr>
<tr>
<td>Stage 0</td>
<td>10 (4.8)</td>
</tr>
<tr>
<td>Stage 1</td>
<td>60 (28.7)</td>
</tr>
<tr>
<td>Stage 2</td>
<td>86 (41.1)</td>
</tr>
<tr>
<td>Stage 3</td>
<td>53 (25.4)</td>
</tr>
<tr>
<td>Time since diagnosis (months), mean (SD)</td>
<td>16.6 (8.0)</td>
</tr>
<tr>
<td>Breast cancer surgery, n (%)</td>
<td>197 (94.3)</td>
</tr>
<tr>
<td>Type of breast cancer surgery, n (%)</td>
<td></td>
</tr>
<tr>
<td>Lumpectomy</td>
<td>103 (49.3)</td>
</tr>
<tr>
<td>Axillary dissection</td>
<td>126 (60.3)</td>
</tr>
<tr>
<td>Mastectomy</td>
<td>97 (46.4)</td>
</tr>
<tr>
<td>Breast reconstruction</td>
<td>25 (12)</td>
</tr>
<tr>
<td>Treatment, n (%)</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>152 (72.7)</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>159 (76.1)</td>
</tr>
<tr>
<td>Targeted therapy</td>
<td>60 (28.7)</td>
</tr>
<tr>
<td>Immunotherapy</td>
<td>8 (3.8)</td>
</tr>
<tr>
<td>Comorbid chronic illness (yes), n (%)</td>
<td>53 (25.4)</td>
</tr>
<tr>
<td>Educational level, n (%)</td>
<td></td>
</tr>
<tr>
<td>Primary education</td>
<td>15 (7.2)</td>
</tr>
<tr>
<td>Secondary education</td>
<td>102 (48.8)</td>
</tr>
<tr>
<td>Tertiary and higher</td>
<td>91 (43.5)</td>
</tr>
<tr>
<td>Refused to answer</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>33 (15.9)</td>
</tr>
<tr>
<td>Married</td>
<td>153 (73.6)</td>
</tr>
<tr>
<td>Divorced or widowed</td>
<td>22 (10.6)</td>
</tr>
<tr>
<td>Monthly household income (HK$), n (%)</td>
<td></td>
</tr>
<tr>
<td>≥10,000</td>
<td>46 (22)</td>
</tr>
<tr>
<td>10,001-30,000</td>
<td>42 (20.5)</td>
</tr>
<tr>
<td>30,001-50,000</td>
<td>43 (20.6)</td>
</tr>
<tr>
<td>&gt;50,000</td>
<td>35 (16.7)</td>
</tr>
<tr>
<td>Refused to answer</td>
<td>42 (20.1)</td>
</tr>
<tr>
<td>Had a religious affiliation, n (%)</td>
<td>99 (47.4)</td>
</tr>
<tr>
<td>Employment status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>72 (34.4)</td>
</tr>
</tbody>
</table>
Intentions to Use Teledelivered SCC Services

Participants’ intentions to use different types of teledelivered SCC services are presented in Table 2. Almost all the teledelivered SCC services listed were accepted by most of the participants. The most accepted teledelivered SCC services in different categories were psychooncology counseling (140/209, 67%), nutrition consultation (165/209, 78.9%), movement and exercise activities (146/209, 69.9%), and patient support groups (131/209, 62.6%).

Table 2. Acceptability of teledelivered supportive cancer care services among breast cancer patients (N=209).

<table>
<thead>
<tr>
<th>Teledelivered supportive cancer care services</th>
<th>Reporting moderate or high intention to use, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Psychosocial care</strong></td>
<td></td>
</tr>
<tr>
<td>Psychotherapy</td>
<td>117 (56)</td>
</tr>
<tr>
<td>Psychological counseling and support</td>
<td>119 (56.9)</td>
</tr>
<tr>
<td>Psychooncology counseling</td>
<td>140 (67)</td>
</tr>
<tr>
<td>Therapist-led group</td>
<td>133 (63.6)</td>
</tr>
<tr>
<td>Cancer prevention and adaption offers for patients and healthy family members</td>
<td>113 (54)</td>
</tr>
<tr>
<td>Family counseling</td>
<td>71 (34)</td>
</tr>
<tr>
<td><strong>Medical consultation</strong></td>
<td></td>
</tr>
<tr>
<td>Cancer helpline</td>
<td>130 (62.2)</td>
</tr>
<tr>
<td>Special medical consultation</td>
<td>132 (63.1)</td>
</tr>
<tr>
<td>To get a second opinion about treatment options</td>
<td>128 (61.2)</td>
</tr>
<tr>
<td>Palliative care consultation</td>
<td>129 (61.7)</td>
</tr>
<tr>
<td>Expert consultation</td>
<td>126 (60.3)</td>
</tr>
<tr>
<td>Nutrition consultation</td>
<td>165 (78.9)</td>
</tr>
<tr>
<td>Complementary and alternative medicine (including traditional Chinese medicine) consultation</td>
<td>139 (66.5)</td>
</tr>
<tr>
<td><strong>Complementary care</strong></td>
<td></td>
</tr>
<tr>
<td>Movement and exercise activities (eg, yoga, qigong, exercises for pain relief)</td>
<td>146 (69.9)</td>
</tr>
<tr>
<td>Creative therapeutic offers (music and art therapy)</td>
<td>105 (50.2)</td>
</tr>
<tr>
<td>Relaxation, breathing, meditation exercise group sessions</td>
<td>121 (57.9)</td>
</tr>
<tr>
<td>Mindfulness exercises</td>
<td>103 (49.2)</td>
</tr>
<tr>
<td>Massage exercises</td>
<td>108 (51.7)</td>
</tr>
<tr>
<td><strong>Peer support groups</strong></td>
<td></td>
</tr>
<tr>
<td>Internet forum with peers</td>
<td>95 (45.5)</td>
</tr>
<tr>
<td>Patient support group</td>
<td>131 (62.6)</td>
</tr>
</tbody>
</table>

Correlations Between Major Variables and Intention to Use Telehealth

The correlation analysis results showed that the participants with a higher education level, prior telehealth experience, and more confidence using technology devices were more likely to report a higher intention to use telehealth (Table 3). Older age was associated with lower intentions to use 3 different types of teledelivered oncology services. Higher levels of performance expectancy, effort expectancy, facilitating conditions, and social influence were associated with higher intentions to use teledelivered oncology services. A higher level of technology anxiety was negatively correlated with intentions to use teledelivered oncology services. Contrary to the hypotheses, fear of COVID-19 was not associated with intentions to use...
teledelivered oncology services (Table 3). Other demographic characteristics (eg, marital status, $P=.82$; cancer stage, $P=.83$; time since diagnosis, $P=.18$; income, $P=.10$) were not correlated with the intention to use telehealth (data not tabulated).

Table 3. Correlations among major independent variables and intentions to use teledelivered supportive cancer care services (N=209).

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Intention to use psychosocial teledelivered supportive care</th>
<th>Intention to use teledelivered medical consultations</th>
<th>Intention to use teledelivered complementary cancer care</th>
<th>Intention to use teledelivered peer support groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$-0.20$</td>
<td>$-0.17$</td>
<td>$-0.10$</td>
<td>$-0.28$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.005</td>
<td>.02</td>
<td>.16</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>2. Education(^b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.22$</td>
<td>$0.16$</td>
<td>$0.24$</td>
<td>$0.32$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.01</td>
<td>.02</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>3. Prior telehealth use(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.29$</td>
<td>$0.24$</td>
<td>$0.22$</td>
<td>$0.32$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>.01</td>
<td>.01</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>4. Confidence using technological devices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.34$</td>
<td>$0.34$</td>
<td>$0.31$</td>
<td>$0.30$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>5. Fear of COVID-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.05$</td>
<td>$0.01$</td>
<td>$-0.05$</td>
<td>$0.03$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.47</td>
<td>.92</td>
<td>.47</td>
<td>.65</td>
</tr>
<tr>
<td>6. Performance expectancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.45$</td>
<td>$0.39$</td>
<td>$0.36$</td>
<td>$0.40$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>7. Effort expectancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.37$</td>
<td>$0.32$</td>
<td>$0.29$</td>
<td>$0.32$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>8. Facilitating conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.41$</td>
<td>$0.34$</td>
<td>$0.30$</td>
<td>$0.43$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>9. Social influence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$0.30$</td>
<td>$0.32$</td>
<td>$0.26$</td>
<td>$0.22$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>.001</td>
</tr>
<tr>
<td>10. Technology anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>$-0.18$</td>
<td>$-0.18$</td>
<td>$-0.14$</td>
<td>$-0.22$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>.009</td>
<td>.01</td>
<td>.04</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

\(^a\leq 55$ years (0); $>55$ years (1).

\(^b\)High school or less (0); at least college (1).

\(^c\)No (0); Yes (1).

Hierarchical Regression Analysis

Given that the independent variables were moderately correlated, the independent variables were checked for multicollinearity in the regression analysis. None of the variables had a variance inflation factor $\geq 5$, which indicated the absence of multicollinearity problems.

In block 1, the background variables explained 16.4%, 14.9%, 13.4%, and 20.2% of the variance in the intentions to use teledelivered psychosocial care, medical consultation, complementary care, and peer support groups, respectively. Specifically, a higher education level was associated with higher intentions to use teledelivered complementary care and peer support groups, and greater confidence with using technological
devices was associated with higher intentions to use all 4 types of teledelivered SCC services. Prior telehealth use was associated with greater intentions to use teledelivered medical consultation and peer support groups (Tables 4 and 5).
Table 4. Hierarchical regression analyses to explain intentions to use telehealth services (N=209).

<table>
<thead>
<tr>
<th>Steps</th>
<th>Intentions to use teledelivered supportive cancer care</th>
<th>Medical consultation&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\Delta R^2$</td>
</tr>
<tr>
<td></td>
<td>Psychosocial care&lt;sup&gt;a&lt;/sup&gt;</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Step 1: Background variables</td>
<td></td>
<td>0.164</td>
</tr>
<tr>
<td>Age&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.06</td>
<td>.67</td>
</tr>
<tr>
<td>Education&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.27</td>
<td>.06</td>
</tr>
<tr>
<td>Prior telehealth use&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.27</td>
<td>.08</td>
</tr>
<tr>
<td>Confidence using technological devices</td>
<td>0.28</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Fear of COVID-19</td>
<td>0.12</td>
<td>.08</td>
</tr>
<tr>
<td>Step 2: Telehealth-related perceptions</td>
<td></td>
<td>0.159</td>
</tr>
<tr>
<td>Age&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.02</td>
<td>.89</td>
</tr>
<tr>
<td>Education&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.12</td>
<td>.39</td>
</tr>
<tr>
<td>Prior telehealth use&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.19</td>
<td>.19</td>
</tr>
<tr>
<td>Confidence using technological devices</td>
<td>0.20</td>
<td>.01</td>
</tr>
<tr>
<td>Fear of COVID-19</td>
<td>0.06</td>
<td>.35</td>
</tr>
<tr>
<td>Performance expectancy</td>
<td>0.34</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Effort expectancy</td>
<td>-0.07</td>
<td>.47</td>
</tr>
<tr>
<td>Facilitating conditions</td>
<td>0.20</td>
<td>.02</td>
</tr>
<tr>
<td>Social influence</td>
<td>0.08</td>
<td>.30</td>
</tr>
<tr>
<td>Technology anxiety</td>
<td>0.08</td>
<td>.28</td>
</tr>
<tr>
<td>Step 3: Interaction terms</td>
<td></td>
<td>0.013</td>
</tr>
<tr>
<td>Age&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.03</td>
<td>.81</td>
</tr>
<tr>
<td>Education&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.13</td>
<td>.36</td>
</tr>
<tr>
<td>Prior telehealth use&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.17</td>
<td>.25</td>
</tr>
<tr>
<td>Confidence using technological devices</td>
<td>0.20</td>
<td>.01</td>
</tr>
<tr>
<td>Fear of COVID-19</td>
<td>0.07</td>
<td>.30</td>
</tr>
<tr>
<td>Performance expectancy</td>
<td>0.30</td>
<td>.01</td>
</tr>
<tr>
<td>Effort expectancy</td>
<td>-0.07</td>
<td>.56</td>
</tr>
<tr>
<td>Facilitating conditions</td>
<td>0.19</td>
<td>.09</td>
</tr>
<tr>
<td>Social influence</td>
<td>0.09</td>
<td>.35</td>
</tr>
<tr>
<td>Technology anxiety</td>
<td>-0.03</td>
<td>.73</td>
</tr>
<tr>
<td>Performance expectancy × education</td>
<td>0.07</td>
<td>.67</td>
</tr>
<tr>
<td>Effort expectancy × education</td>
<td>0.02</td>
<td>.90</td>
</tr>
<tr>
<td>Facilitating conditions × education</td>
<td>0.03</td>
<td>.85</td>
</tr>
<tr>
<td>Social influence × education</td>
<td>-0.03</td>
<td>.82</td>
</tr>
<tr>
<td>Technology anxiety × education</td>
<td>0.26</td>
<td>.07</td>
</tr>
</tbody>
</table>

<sup>a</sup>Total $R^2$: 0.336.
<sup>b</sup>Total $R^2$: 0.281.
<sup>c</sup>≤55 years (0); >55 years (1).
<sup>d</sup>High school or less (0); at least college (1).
<sup>e</sup>No (0); Yes (1).
Table 5. Hierarchical regression analyses to explain intentions to use telehealth services (N=209).

<table>
<thead>
<tr>
<th>Step</th>
<th>Intentions to use tele-delivered supportive cancer care</th>
<th>Peer support groups&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Complementary care&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ΔR²</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>P value</td>
</tr>
<tr>
<td>Step 1: Background variables</td>
<td>0.134</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.17</td>
<td>.26</td>
</tr>
<tr>
<td>Education&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.36</td>
<td>.01</td>
</tr>
<tr>
<td>Prior telehealth use&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.21</td>
<td>.19</td>
</tr>
<tr>
<td>Confidence using technological devices</td>
<td>0.27</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Fear of COVID-19</td>
<td>0.01</td>
<td>.87</td>
</tr>
</tbody>
</table>

Step 2: Telehealth-related perceptions 0.096 <.001 0.122 <.001

| Age<sup>c</sup> | 0.18 | .23 | −0.19 | .17 |
| Education<sup>d</sup> | 0.28 | .05 | 0.32 | .02 |
| Prior telehealth use<sup>e</sup> | 0.17 | .26 | 0.30 | .04 |
| Confidence using technological devices | 0.23 | .01 | 0.08 | .28 |
| Fear of COVID-19 | −0.05 | .50 | 0.08 | .22 |
| Performance expectancy | 0.25 | .002 | 0.30 | <.001 |
| Effort expectancy | −0.05 | .64 | −0.17 | .08 |
| Facilitating conditions | 0.07 | .44 | 0.26 | .002 |
| Social influence | 0.13 | .09 | 0.03 | .70 |
| Technology anxiety | 0.06 | .41 | −0.02 | .77 |

Step 3: Interaction terms 0.031 .16 0.029 .13

| Age<sup>c</sup> | 0.20 | .17 | −0.21 | .13 |
| Education<sup>d</sup> | 0.34 | .02 | 0.32 | .02 |
| Prior telehealth use<sup>e</sup> | 0.13 | .40 | 0.31 | .03 |
| Confidence using technological devices | 0.22 | .01 | 0.07 | .38 |
| Fear of COVID-19 | −0.02 | .79 | 0.08 | .24 |
| Performance expectancy | 0.12 | .30 | 0.25 | .02 |
| Effort expectancy | −0.00 | .10 | −0.05 | .67 |
| Facilitating conditions | 0.17 | .17 | 0.11 | .32 |
| Social influence | 0.15 | .14 | 0.05 | .62 |
| Technology anxiety | 0.01 | .94 | −0.11 | .23 |
| Performance expectancy × education | 0.34 | .04 | 0.05 | .77 |
| Effort expectancy × education | −0.10 | .63 | −0.29 | .13 |
| Facilitating conditions × education | −0.27 | .13 | 0.36 | .03 |
| Social influence × education | −0.09 | .56 | 0.01 | .92 |
| Technology anxiety × education | 0.09 | .56 | 0.22 | .12 |

<sup>a</sup>Total R²: 0.261.
<sup>b</sup>Total R²: 0.353.
<sup>c</sup>≤55 years (0); >55 years (1).
<sup>d</sup>High school or less (0); at least college (1).
<sup>e</sup>No (0); Yes (1).
In block 2, telehealth-related perceptions explained an additional 15.9%, 12.6%, 9.6%, and 12.2% of the variance in intentions to use teledelivered psychosocial care, medical consultation, complementary care, and peer support groups, respectively. Specifically, performance expectancy was associated with intentions to use all 4 types of teledelivered SCC services (Tables 4 and 5). More facilitating conditions were associated with higher intentions to use teledelivered psychosocial care and peer support groups. Greater social influence was associated with higher intentions to use teledelivered medical consultation ($\beta=0.16, P=.03$; Tables 4 and 5).

In block 3, 5 interaction terms between education level and telehealth-related perceptions were entered; 2 significant interactions emerged. Specifically, there was an interaction between education level and performance expectancy when explaining the intention to use teledelivered complementary care ($\beta=0.34, P=.04$). In addition, there was an interaction between education level and facilitating conditions when explaining the intention to use teledelivered peer support groups ($\beta=0.36, P=.03$). Simple slopes analysis results indicated that the association between performance expectancy and intention to use teledelivered complementary care was only significant among those with a higher education level ($\beta=0.46, P<.001$) but not among those with a lower education level ($\beta=0.12, P=.30$; Figure 2). Similarly, the association between social influence and intention to use teledelivered peer support groups was only significant among those with a higher education level ($\beta=0.48, P<.001$) but not among those with a lower education level ($\beta=0.11, P=.32$; Figure 3). Overall, the models explained 26.1% to 35.3% of the variance in the intentions to use different types of teledelivered SCC services (Tables 4 and 5).

Figure 2. Relationship between performance expectancy and intention to use teledelivered complementary care by education level.
Discussion

Principal Findings

This study examined how sociodemographic and clinical factors and telehealth-related perceptions contributed to the intentions to use telehealth for SCC among BCS in Hong Kong during the COVID-19 pandemic. It is noteworthy that most of the participants reported moderate-to-high intentions to use different types of teledelivered SCC services. The most accepted teledelivered SCC services in different categories were psychooncology counseling (67%), nutrition consultation (78.9%), movement and exercise activities (69.9%), and patient support groups (62.6%). We found that greater confidence in telehealth use, performance expectancy (believing telehealth helps with daily tasks), social influence (important others encouraging telehealth use), and facilitating conditions (having resources for telehealth use) were associated with higher intentions to use teledelivered SCC. Our findings were comparable to those of a study in Singapore amid the COVID-19 pandemic [38] that showed that general acceptance of telemedicine by patients with cancer was around 60%. Perceptions that telemedicine could improve health care access and the availability of necessary resources for telemedicine were associated with higher acceptance among those patients [38].

Sociodemographic Factors, Fear of COVID-19, and Intention to Use Teledelivered SCC

In our regression analyses, education level, prior telehealth use, and confidence using technological devices were associated with the use of telehealth services. Our findings were consistent with findings from patient populations in Western countries supporting that people with mobile device access, who were confident using technological devices, and with prior telehealth experience were more likely to use teledelivered SCC [38-40]. The facilitating roles of those variables seem to be culturally and geographically universal. To increase patients’ intentions to use teledelivered SCC, it might be important to provide education and training on how to use technology and telehealth services, which could help increase confidence with using these tools and make it easier for patients to access care.

Consistent with a population-based study in the United States during the COVID-19 pandemic [41], household income was not a significant contributor to intentions to use teledelivered SCC in our study. However, the findings should be interpreted with caution, as a high proportion of participants (20.1%) refused to report their household income. Household income has been associated with other important sociodemographic factors (eg, education, ownership of mobile devices, internet access) that were associated with cancer survivors’ intentions to adopt telehealth before and during the COVID-19 pandemic [42,43]. Given that 97.6% of our participants possessed a mobile phone with internet access, the unique contribution of household income on intention to use telehealth might become less apparent.

Despite a significant bivariate correlation between age and intention to use teledelivered SCC, age did not emerge as a significant contributor in the regression analyses beyond the influence of other potential contributors. These findings imply that other individual characteristics (eg, confidence using...

Figure 3. Relationship between facilitating conditions and intention to use teledelivered peer support groups.

\[ \beta = 0.48, \ p < 0.01 \]

- Low education level
- High education level

\[ \beta = 0.11, \ p = 0.32 \]
technological devices) played a stronger role in the intentions of BCS to use teledelivered SCC. Moreover, it is important to note that Hong Kong has a very high internet coverage rate at the household level (96.1%) and a very high smartphone ownership rate (99.8% and 90.7% among individuals aged 45-64 years or ≥65 years, respectively) [44], which could influence the acceptability of and perceptions toward telehealth services. The generalizability of our findings to other countries with different internet use patterns should also be interpreted with caution [45].

Fear of COVID-19 did not emerge as a significant contributor to the intention to use teledelivered SCC services in our sample, which was contrary to the findings of prior studies in the United States [19] and Germany [46]. However, our findings seem to be in line with those of An and colleagues [47] who showed that anxiety about COVID-19 was not associated with telehealth acceptance among individuals with chronic disease in South Korea. A potential reason for the discrepancies in the findings could be related to the focus of the measurements. The Fear of COVID-19 scale used in this study primarily measures participants’ affective responses and anxiety symptoms toward cues related to COVID-19, but it might not capture individuals’ perceptions of the threat of contracting COVID-19 at different occasions (eg, hospital and clinic settings, crowded places). Such concerns have been reflected in studies among BCS [48]. Future studies might elucidate how patients’ specific COVID-19 worries and concerns contribute to their intentions to use telehealth services.

**Teledelivered SCC: Education Level as a Moderator**

We found that education level moderated the interaction between (1) between performance expectancy and intention to use teledelivered complementary care and (2) facilitating conditions and intention to use teledelivered peer support groups. From the perspective of the UTAUT model, performance expectancy (ie, degree to which the individual believes that using the technology will help them better cope with daily life or be more effective) was found to be associated with higher intentions to use teledelivered complementary care (including creative therapies, relaxation, and mindfulness exercises) only among those with a higher education level. It is also noteworthy that similar patterns of findings were also apparent in other aspects of technology use. Education level moderated the positive associations between technology use perceptions (performance expectancy, facilitating conditions) and people’s intentions to use mobile banking services in Jordan [23].

Our findings suggested that just highlighting performance expectancy might not be sufficient to significantly increase intentions to use teledelivered complementary care among those with a lower education level. A basic understanding of those complementary care options might be important. Given that those with higher levels of education may be more likely to have better awareness of the potential benefits of those complementary therapies for oncology care [53], the facilitating role of performance expectancy in the intention of BCS to use teledelivered complementary care could be strengthened by a higher education level.

Similarly, we found that facilitating conditions were associated with higher intentions to use teledelivered peer support groups only among those with a higher education level. Facilitating conditions refer to people’s perceptions about whether the necessary resources and support are available to use the technology effectively. It is important to note that peer support groups generally involve mutual interactions and sharing with other cancer survivors, which could also be subject to challenges such as confrontation involving others’ suffering, divergent groups (but not other types of SCC). It is noteworthy that psychological care and peer support group services are not commonly utilized among local BCS [25]. The dynamics in psychological counseling and peer support groups involve more disclosure of personal challenges and distress, which might be incongruent with the cultural preference of not bringing up negative emotions to maintain social harmony [33]. It might be possible for local BCS to believe that they need a certain level of knowledge and informational resources (facilitating conditions) to understand what to expect in teledelivered psychosocial care and peer support groups before enrolling in those services.
information needs, conflicts in group dynamics, and challenges with sustainability [54]. Individuals with higher levels of education may be more comfortable using teledelivered services to interact with other patients with similar (stressful) experiences plus have more resources to deal with the potentially negative experiences in the support group context (eg, worsened health of peers in the group, appraising information about their illness, and therapy options shared in the support groups). These reasons might explain why the facilitating role of facilitating conditions in the intention of BCS to use teledelivered peer support group was only apparent among those with a higher education level.

Limitations
This study was subject to several limitations. First, this study used a cross-sectional design, which might not highlight the causal relationship among the variables. Cancer survivors’ expectations and motivations for teledelivered cancer care may also change over time. Future studies could use longitudinal designs to better understand the temporal relationships among the variables and their future use of teledelivered care services. Second, to allow more systematic recruitment of recently diagnosed BCS (since the COVID-19 outbreak in Hong Kong), we recruited BCS through local cancer registries. Even though the HKBCR is the most comprehensive registry for BCS in Hong Kong, it is noteworthy that not everyone in the total BCS population was covered due to the HKBCR’s voluntary enrollment system. Based on the Hong Kong Cancer Registry data [55] and HKBCR [56] for individuals with BCS aged 18 years to 70 years, the age group distributions were as follows: 40% (<50 years), 33% (50-59 years), 27% (60-70 years). Similarly, in our sample, the age group distributions were as follows: 39.2% (<50 years), 30.1% (51-60 years), and 29.7% (≥61 years). Our sample was highly comparable in terms of the age distribution of the local BCS. However, the generalizability of the findings to BCS in other regions or countries with different health care systems and to survivors of other cancer types might be limited. Third, the studied variables only explained a moderate proportion of variance in the intentions of BCS to use teledelivered SCC. Other factors might be at play. Research has found that other telehealth-related perceptions (eg, privacy concerns), the specific characteristics of different teledelivered services (eg, expected durations and schedules of the services, the necessity to use cameras for the services, group- and individual-based delivery), and contextual factors (eg, severity of the pandemic situation, availability of specific types of teledelivered care services) could be important determinants for those intentions [18,57]. Consideration of those variables might further improve the explanatory power of the regression model.

Implications
The COVID-19 pandemic impacted cancer service utilization among cancer patients worldwide. Telehealth can be a new service model for SCC services, especially after the experience of the COVID-19 pandemic. The use of telehealth for SCC not only provides flexibility for services in hospitals and cancer clinics but also potentially improves cancer survivors’ well-being. Recent reviews and trials have found that teledelivered interventions facilitate positive physical and psychological health impacts on cancer survivors [58-60]. Therefore, identifying the potential determinants for people’s intentions to use telehealth for SCC could facilitate the proposal of novel service models.

This was one of the first attempts to examine how telehealth-related perceptions, sociodemographic and clinical characteristics, and cancer service utilization experiences during COVID-19 contributed to the intention of BCS to use telehealth for SCC during the COVID-19 pandemic in Hong Kong. It is essential for health care providers to be knowledgeable about specific factors facilitating the intention to use telehealth, so that patients’ needs and cancer care preferences can be met, especially for the response to a potential pandemic of an emerging infectious disease in the future.

Researchers have started to advocate for a patient-centered approach to address patients’ facilitators and barriers to using telehealth. By fitting telehealth into the overall patient journey and treatment plan and applying inclusive design principles, the needs of the most vulnerable populations who may not be engaging with telehealth owing to their age, education level, socioeconomic status, technology skills, and experiences could be better addressed [40]. Our findings imply that enhancing BCS’ skills for using telehealth, improving BCS’ and their important others’ perceived benefits of telehealth, and providing assistance for telehealth use could increase BCS’ intentions to use teledelivered SCC. For intentions to use specific types of SCC (eg, complementary care and peer support groups), addressing relevant factors (performance expectancy, facilitating conditions) might be particularly beneficial for those with a higher education level.

Acknowledgments
This study was supported by the Health and Medical Research Fund (reference number 18190061; principal investigator: NCYY).

Data Availability
Research data will be shared upon request.

Authors’ Contributions
NCYY contributed to the supervision and conceptualization of the study and acquired funding. NCYY and STYL were involved in the data curation, conducted the data analysis, and wrote the original manuscript draft. PSYC provided advice on the
implementation of the study using the Hong Kong Breast Cancer Registry. NCYY, STYL, EYYC, CC, WWSM, JYMS, and PSYC contributed to the methodology, survey design, and review and editing of the manuscript.

Conflicts of Interest
None declared.

References


47. An M, You S, Park R, Lee S. Using an extended technology acceptance model to understand the factors influencing telehealth utilization after flattening the COVID-19 curve in South Korea: cross-sectional survey study. JMIR Med Inform 2021 Jan 08;9(1):e25435 [FREE Full text] [doi: 10.2196/25435] [Medline: 33395397]


Abbreviations

BCS: breast cancer survivors
HKBCR: Hong Kong Breast Cancer Registry
SCC: supportive cancer care
UTAUT: unified theory of acceptance and use of technology

Edited by T de Azevedo Cardoso; submitted 20.07.23; peer-reviewed by M Coccia, M Rogers, K Bera; comments to author 29.02.24; revised version received 22.03.24; accepted 16.04.24; published 27.06.24.
Please cite as:
Yeung N CY, Lau STY, Mak WWS, Cheng C, Chan EYY, Siu JYM, Cheung PSY
Applying the Unified Theory of Acceptance and Use of Technology to Identify Factors Associated With Intention to Use Tele-delivered Supportive Care Among Recently Diagnosed Breast Cancer Survivors During COVID-19 in Hong Kong: Cross-Sectional Survey
JMIR Cancer 2024;10:e51072
URL: https://cancer.jmir.org/2024/1/e51072
doi:10.2196/51072
PMID:
Usability and Preliminary Efficacy of an Adaptive Supportive Care System for Patients With Cancer: Pilot Randomized Controlled Trial

Sharon H Baik1, PhD; Karen Clark1, MS; Marisol Sanchez1, BS; Matthew Loscalzo1, MSW; Ashley Celis1, BS; Marianne Razavi1, PhD; Dershung Yang2, PhD; William Dale1, MD, PhD; Niina Haas2, MA

1Department of Supportive Care Medicine, City of Hope, Duarte, CA, United States
2BrightOutcome, Buffalo Grove, IL, United States

Corresponding Author:
Sharon H Baik, PhD
Department of Supportive Care Medicine
City of Hope
1500 East Duarte Road
Duarte, CA, 91010
United States
Phone: 1 626 256 4673
Email: shbaik@coh.org

Abstract

Background: Using an iterative user-centered design process, our team developed a patient-centered adaptive supportive care system, PatientCareAnywhere, that provides comprehensive biopsychosocial screening and supportive cancer care to patients across the continuum of care adaptively. The overarching goal of PatientCareAnywhere is to improve health-related quality of life (HRQOL) and self-efficacy of patients with cancer by empowering them with self-management skills and bringing cancer care support directly to them at home. Such support is adaptive to the patient’s needs and health status and coordinated across multiple sources in the forms of referrals, education, engagement of community resources, and secure social communication.

Objective: This study aims to assess the usability of the new web-based PatientCareAnywhere system and examine the preliminary efficacy of PatientCareAnywhere to improve patient-reported outcomes compared with usual care.

Methods: For phase 1, usability testing participants included patients with cancer (n=4) and caregivers (n=7) who evaluated the software prototype and provided qualitative (eg, interviews) and quantitative (eg, System Usability Scale) feedback. For phase 2, participants in the 3-month pilot randomized controlled trial were randomized to receive the PatientCareAnywhere intervention (n=36) or usual care control condition (n=36). HRQOL and cancer-relevant self-efficacy were assessed at baseline (preintervention assessment) and 12 weeks from baseline (postintervention assessment); mean differences between pre- and postintervention scores were compared between the 2 groups.

Results: Participants were highly satisfied with the prototype and reported above-average acceptable usability, with a mean System Usability Scale score of 84.09 (SD 10.02). Qualitative data supported the overall usability and perceived usefulness of the intervention, with a few design features (eg, “help request” function) added based on participant feedback. With regard to the randomized controlled trial, patients in the intervention group reported significant improvements in HRQOL from pre- to postintervention scores (mean difference 6.08, SD 15.26) compared with the control group (mean difference −2.95, SD 10.63; P=0.01). In contrast, there was no significant between-group difference in self-efficacy (P=.09).

Conclusions: Overall, PatientCareAnywhere represents a user-friendly, functional, and acceptable supportive care intervention with preliminary efficacy to improve HRQOL among patients diagnosed with cancer. Future studies are needed to further establish the efficacy of PatientCareAnywhere as well as explore strategies to enhance user engagement and investigate the optimal intensity, frequency, and use of the intervention to improve patient outcomes.

Trial Registration: ClinicalTrials.gov NCT02408406; https://clinicaltrials.gov/study/NCT02408406

doi:10.2196/49703
KEYWORDS
cancer; distress screening; eHealth; supportive care; mobile phone

Introduction

Background

One-third [1,2] to half [3-5] of patients with cancer report psychological distress. Common causes of distress include fatigue, pain, worry about the future, finances, and the side effects of treatment [6-8]. Supportive care is a complex specialty that encompasses an array of multidisciplinary services addressing a variety of biopsychosocial concerns and needs. The 2008 Institute of Medicine report, Cancer Care for the Whole Patient [6], lists the main supportive care services as “information about illness, treatments, health, and services; help in coping with emotions accompanying illness and treatment; help in managing illness; assistance in changing behaviors to minimize impact of disease; material and logistical resources, such as transportation; help in managing disruptions in work, school, and family life; and financial advice and/or assistance.”

In addition to these formal sources of supportive care, the report stressed that informal sources, such as family and friends, are also key providers of supportive care. At the heart of a successful supportive care practice is comprehensive biopsychosocial screening, covering multiple domains including physical symptoms, psychosocial issues, and practical concerns. Effective biopsychosocial screening integrated with triage, referrals, patient and caregiver education, and follow-up services promotes successful whole patient-centered care across the cancer treatment trajectory. Studies have demonstrated that adequate integration of biopsychosocial screening with supportive care results in better patient outcomes [9-18], better patient-provider communication [9,12,15,19-25], higher patient satisfaction [12,20,22-24], detection of unrecognized problems [10,12,15,21,23-25], improved referrals [11,25-29], and better health service use and lower costs [30-35].

Recognizing the importance of distress management, the National Comprehensive Cancer Network (NCCN) recommends distress screening for all patients with cancer to address problems before a crisis develops and necessitates higher levels of intervention, with guidelines in place since 1999 [36]. Unfortunately, a serious gap remains between the screening services that are needed and those provided today [6,37,38]. In 2018 a survey to NCCN member institutions, 87% (20/23) of institutions reported conducting routine screening for distress as per the guidelines, but only 26% (6/23) conducted interviews of all patients and 57% (13/23) screened outpatients only [39]. Compared with the 2012 survey [37], the percentage of institutions conducting screening of all patients decreased from 30% to 26% and the percentage of institutions screening outpatients only increased from 50% to 57% over a 6-year span. Most institutions administered screening via paper and pencil (12/23, 52%) or electronically (12/23, 52%), while 30% conducted interviews (6 in person and 1 via telephone). In addition, only 7 institutions reported automatic triage based on computer-generated results, whereas 14 institutions required clinical staff to manually review the screening results to generate referrals.

Furthermore, there is often a large gap between the onset of patients’ distress and the communication about it to their health care team, especially when these problems and symptoms occur outside of the clinical environment. This disconnect is exacerbated by the lack of uniform systems to document problems and communications between the health care professionals themselves. In addition, the absence of systematic criteria-based identifiers for referring patients to suitable consultation services and resources results in important clinical information not being communicated promptly to the appropriate professionals. Electronic methods for distress screening, including automated touch screen technologies and web-based assessments, have been recommended as they can be helpful with systematically identifying, tracking, and managing sources of distress [40]. Over the past decade, technology and eHealth interventions have increasingly been used in the delivery of patient-centered cancer care [41-43]. A recent systematic review of technology-based supportive care interventions for patients with cancer demonstrated significant effects on health-related quality of life (HRQOL), cancer-related symptoms, levels of fatigue and pain, depression, and functional capacity [44]. A meta-analysis was precluded due to heterogeneity in intervention design and features (eg, duration, frequency, and use of technology) and outcome measures.

To address this pressing gap in supportive cancer care, City of Hope in partnership with BrightOutcome, a health care technology company, developed a technology-based, patient-centered adaptive supportive care system for patients newly diagnosed with cancer (named PatientCareAnywhere) using an iterative user-centered design process. PatientCareAnywhere was derived from two existing systems: (1) SupportScreen from City of Hope [45], a clinic-based biopsychosocial screening tool that connects new patients with individualized educational and professional symptom triage support based on self-reported distress; and (2) MyCaringCircle from BrightOutcome, a home- and community-based patient portal solution that offers self-reported symptom assessment, individualized education content delivery, facilitation of remote medical care, and coordination of support from the patient’s friends and family and from community resources. While SupportScreen excels in the provision of a broad range of biopsychosocial screenings, facilitation of referrals, and integration of electronic health records (EHRs), MyCaringCircle’s strengths are its focus on symptom assessment via its access to a large library of validated measures and its facilitation of social support outside the clinical environment involving community resources.

Objective

This study includes 2 phases. In phase 1, with the software prototype, we conducted usability tests, which are an integral part of the user-centered design process and help ensure the intervention meets users’ expectations and functions as intended. In phase 2, to evaluate the preliminary efficacy of PatientCareAnywhere compared with usual care (control condition), we conducted a pilot randomized controlled trial
(RCT) evaluating changes in self-reported patient outcomes, including HRQOL and self-efficacy, from baseline to postintervention assessment. We hypothesized that PatientCareAnywhere would result in significant improvements in HRQOL and patient self-efficacy compared with usual care among patients newly diagnosed with cancer.

**Methods**

**PatientCareAnywhere**

**Overview**

City of Hope, in partnership with BrightOutcome, a health care technology company, developed a patient-centered adaptive supportive care system (PatientCareAnywhere) to improve patient outcomes for patients with cancer while reducing health care costs. This project was funded by the National Cancer Institute via a Small Business Innovation Research Fast-Track grant (R44CA192588). PatientCareAnywhere is a patient empowerment solution that promotes internal resilience, self-efficacy, and independence. The key features of PatientCareAnywhere include (1) multilevel and adaptive biopsychosocial screening covering a comprehensive set of supportive cancer care domains (eg, emotional, physical, practical, and social) without overburdening patients with long static questionnaires; (2) automatic alert messages for abnormal screening results to clinical team; (3) specialist referrals and community support resources based on screening results; (4) individualized patient education contents based on screening results; (5) social media support for engagement of caregivers, family, friends, and community resources; (6) optimized display for different devices (eg, smartphones and tablets); and (7) EHR integration. The PatientCareAnywhere experience begins with an initial comprehensive biopsychosocial assessment covering physical symptoms (eg, pain), psychosocial issues (eg, anxiety), and practical concerns (eg, finances). Table 1 provides a list of biopsychosocial screening topics and designated care professionals for follow-up. The assessments start with first-level questions, which, when a patient’s response exceeds a pre-established threshold, will trigger additional follow-up questions to gain further insights into the patient’s needs and concerns. Additionally, alert messages are generated for the clinical and support care teams.

These self-reported needs and the individual’s disease and treatment stages form the basis for PatientCareAnywhere to offer responsive supportive care in terms of individualized patient education content, triage to specialists, and referrals to community resources. Cancer-specific content (eg, information about breast, lung, or prostate cancer and its treatment) and generic content (eg, emotional distress) was adapted from public domain sources, such as the National Cancer Institute, American Cancer Society, and NCCN, and from materials developed by the Division of Patient and Family Community Education at City of Hope. We also collected contact information for supportive care services provided by City of Hope and local community resources, which were recommended to patients based on their self-reported symptoms and needs. The system was designed to be used by patients, friends and families, health care professionals, and community resources. With PatientCareAnywhere, patients are at the center of the “circle of care,” receiving support from multiple clinical, social, and community sources and across the continuum of care, from diagnosis to treatment to survivorship and end-of-life care. In addition, PatientCareAnywhere provides a communication platform to allow caregivers, family members, and friends to interact directly with the patient through the system. As a security feature, patients have complete control over who is included in their care circles and how much communication or information is shared with each person invited. In particular, caregivers are granted full access to patient medical records and can obtain information about the patient’s current medications, laboratories and tests, vitals, biopsychosocial screenings, and symptom histories as well as keep track of medical appointments on PatientCareAnywhere, while noncaregivers have limited access. The main components of PatientCareAnywhere are listed in Multimedia Appendix 1, and screenshots of PatientCareAnywhere are included in Multimedia Appendix 2.
### Table 1. Biopsychosocial screening items and associated referrals.

<table>
<thead>
<tr>
<th>Biopsychosocial screening items</th>
<th>Primary follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social and practical needs</strong></td>
<td></td>
</tr>
<tr>
<td>Ability to have children</td>
<td>Physician</td>
</tr>
<tr>
<td>Communication about medical care</td>
<td>Pharmacy clinical manager, physician, social worker</td>
</tr>
<tr>
<td>Finding local support resources</td>
<td>Social worker</td>
</tr>
<tr>
<td>Finding reliable medical information</td>
<td>Cancer information resource nurse, nurse</td>
</tr>
<tr>
<td>Health insurance</td>
<td>Financial counselor</td>
</tr>
<tr>
<td>Help with home or medical care</td>
<td>Patient navigator, resource coordinator, social worker</td>
</tr>
<tr>
<td>Hospice service</td>
<td>Physician, nurse practitioner or physician extender</td>
</tr>
<tr>
<td>Personal finances</td>
<td>Social worker</td>
</tr>
<tr>
<td>Physical appearance</td>
<td>Positive image center, social worker</td>
</tr>
<tr>
<td>Social support</td>
<td>Social worker</td>
</tr>
<tr>
<td>Spiritual or religious concerns</td>
<td>Chaplin</td>
</tr>
<tr>
<td>Worries about the future</td>
<td>Social worker</td>
</tr>
<tr>
<td><strong>Physical and emotional well-being</strong></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>Social worker</td>
</tr>
<tr>
<td>Appetite loss</td>
<td>Nurse</td>
</tr>
<tr>
<td>Bladder control</td>
<td>Nurse</td>
</tr>
<tr>
<td>Breathing difficulties</td>
<td>Nurse</td>
</tr>
<tr>
<td>Bowel control</td>
<td>Nurse</td>
</tr>
<tr>
<td>Cognitive issues</td>
<td>Physician</td>
</tr>
<tr>
<td>Constipation</td>
<td>Nurse</td>
</tr>
<tr>
<td>Depression</td>
<td>Social worker</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>Nurse</td>
</tr>
<tr>
<td>Eating, chewing, or swallowing difficulties</td>
<td>Clinical nutritionist, nurse</td>
</tr>
<tr>
<td>Fatigue</td>
<td>Nurse practitioner or physician extender</td>
</tr>
<tr>
<td>Fever</td>
<td>Nurse</td>
</tr>
<tr>
<td>Mobility or physical issues</td>
<td>Nurse</td>
</tr>
<tr>
<td>Mouth sores</td>
<td>Nurse</td>
</tr>
<tr>
<td>Nausea or vomiting</td>
<td>Nurse</td>
</tr>
<tr>
<td>Numbness</td>
<td>Nurse</td>
</tr>
<tr>
<td>Pain</td>
<td>Physician</td>
</tr>
<tr>
<td>Sexual issues</td>
<td>Nurse</td>
</tr>
<tr>
<td>Skin rash</td>
<td>Nurse</td>
</tr>
<tr>
<td>Sleep issues</td>
<td>Nurse practitioner or physician extender</td>
</tr>
<tr>
<td>Swelling</td>
<td>Nurse</td>
</tr>
<tr>
<td>Weight change</td>
<td>Clinical nutritionist, nurse</td>
</tr>
</tbody>
</table>

### Prototype Design and Development

#### User-Centered Design

The PatientCareAnywhere prototype was developed using an iterative user-centered design approach, in which targeted end users (patients with cancer) and other key stakeholders (eg, caregivers and health care professionals) were involved in the design and development process to ensure the intervention aligns with the needs and preferences of patients newly diagnosed with cancer (target population). Research has shown that involving stakeholders throughout intervention development and evaluation is essential to increasing user acceptance and intervention effectiveness [46,47]. In addition, an expert panel with expertise in the fields of supportive care, oncology, nursing,
mental health, outcome research, palliative care, and patient education was assembled to provide continuous guidance and consultation on our prototype design and evaluation efforts.

**Stakeholder Input**

Feedback from patients with cancer and caregivers was largely unanimous in agreement with the idea of a system such as PatientCareAnywhere and the functions that they would like to see implemented. Notably, they all expressed strong support for social networking functions, the ability to keep track of appointments and medical records, access to tailored recommendations for educational support materials and local events and support groups, and the ability to report symptoms at any time that would send alerts to their care team. Patients expressed a strong interest in being able to connect with other patients with cancer who are going through or have been through the same experiences. Caregivers expressed support for the ability to connect with other caregivers to build a support network of others who are also going through the same caregiver experiences. Finally, patients and caregivers felt that the ability to create “help requests” that they could share with their network would make the logistics involved with having cancer and caring for someone with cancer a lot easier. All participants felt that they would like to use PatientCareAnywhere when it was available and that it would be a great resource for others in the same position. The only barriers that these focus groups identified involved possibly leaving out those who are not as technology savvy. However, each group concluded that most people have someone around who is able to help them with the technology.

Feedback from the expert panel highlighted a number of features that they wanted to see implemented in PatientCareAnywhere and the barriers that they foresaw in using PatientCareAnywhere. Overall, the expert panel liked the idea of a system such as PatientCareAnywhere for clinic use and clinic-based research. All members of the expert panel immediately recognized the benefits of having features such as social networking, tailored educational materials, event recommendations, and symptom reporting and management for patients with cancer and felt that PatientCareAnywhere would enable them to provide better care to their patients. The expert panel members also wanted to have the information from PatientCareAnywhere to be integrated into the EHR or have the 2 systems “speak” to each other so that they only had to enter information into 1 system, and it would automatically populate into both systems. The members also wanted to have additional clinical research features available as part of the initial biopsychosocial screening tool to deliver specialized questionnaires to the patients who are part of different research studies at City of Hope.

**Ethical Considerations**

All study procedures and assessments were reviewed and approved by the City of Hope Institutional Review Board before participant enrollment (institutional review board #15025). Written informed consent was provided by all study participants recruited for the usability testing (phase 1) and pilot RCT (phase 2), and all participants were provided the ability to opt out of the study at any time. To ensure participant privacy and confidentiality, study data were deidentified using participant ID numbers. The mapping between participant IDs and actual participant identities was maintained by the City of Hope research team in a password-protected electronic file. Each study participant was given a unique participant login ID to access the prototype system, which also enabled researchers to retrieve information related to a specific participant. Usability testing participants (phase 1) were compensated US $50 for their time. Pilot RCT participants (phase 2) received a US $100 stipend as compensation for the time spent in the study.

**Phase 1: Usability Testing**

**Overview**

We conducted 2 types of usability testing to evaluate the usability, usefulness, and acceptability of the prototype system. The first usability test was “design oriented” and conducted after wireframes (schematics showing information elements and page flows) were produced. This allowed us to resolve initial design issues before significant development efforts took place. Once most of the development work was completed, we then conducted “metric-oriented” usability tests to formally evaluate the usability of the PatientCareAnywhere using quantitative assessments.

**Study Participants and Design**

To be eligible to participate in usability testing, patients were required to be (1) aged ≥21 years, (2) diagnosed with any cancer, (3) currently receiving any type of cancer treatment, (4) treated on an outpatient status (participation was suspended during hospitalization), (5) fluent in English, and (6) able to access the internet at home. Caregivers, friends, and family members of patients with cancer were also eligible to participate in the study. Those with evidence of cognitive or psychological impairment as well as prisoners and pregnant women were ineligible. Participants were also excluded if they were currently participating in another psychosocial study.

All patients with cancer and caregivers were recruited from City of Hope, a National Cancer Institute–designated comprehensive cancer center in Duarte, California, via physician referrals, subject recruitment flyers, and a touch screen biopsychosocial screening system (SupportScreen [45]), which included a question about participating in this study. Trained research assistants approached potentially eligible patients and discussed study participation either in person during an already scheduled clinic visit or via telephone. Interested patients were then screened for eligibility criteria, and those eligible wishing to enroll provided written informed consent. All participants consented before study participation and were enrolled between March and April 2016.

Each participant completed a 60-minute one-on-one usability testing session, in which they completed specific tasks using the prototype, and an observer recorded how the tasks were completed (or failed). Participants were asked to talk aloud as they performed the tasks. After completing all assigned tasks, participants for the second usability test also completed self-report measures to evaluate perceived usability, usefulness, and acceptability of the PatientCareAnywhere prototype.
Usability testing sessions were audio recorded using encrypted audio recorders and professionally transcribed. The audio files were transmitted via secure protocols to an encrypted project folder on a secure file server at City of Hope. The original audio files were permanently deleted from the audio recorders once uploaded to the file server.

**Measures**

**Usability**

The System Usability Scale (SUS) is a validated and widely used 10-item usability measure [48]. Participants’ scores for each item are added together and then multiplied by 2.5 to convert the original scores of 0 to 40 to 0 to 100, with higher scores indicating higher usability [48]. Overall SUS scores ≥70.0 are considered above average in terms of acceptable usability [49,50].

**Usefulness**

Participants also completed a 35-item Usefulness Questionnaire, which was developed specifically for PatientCareAnywhere and includes statements assessing the usefulness and design features of the system. Participants rated their level of agreement with each statement using a 5-point Likert scale ranging from 1=“strongly disagree” to 5=“strongly agree.”

**Data Analytic Plan**

Data were collected and analyzed using SPSS (IBM Corp). Descriptive statistics (eg, means, frequencies, percentages) were used to characterize the sociodemographic and clinical characteristics of the study sample. Summary statistics were used to describe the usability outcomes, including overall SUS scores and perceived acceptability and usefulness ratings.

**Phase 2: Pilot RCT**

**Study Participants and Design**

Textbox 1 shows the inclusion and exclusion criteria for the pilot RCT.

<table>
<thead>
<tr>
<th><strong>Inclusion criteria</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>At least 21 years of age</td>
</tr>
<tr>
<td>Diagnosis of breast, lung, or prostate cancer at any stage</td>
</tr>
<tr>
<td>Currently being treated on an outpatient basis</td>
</tr>
<tr>
<td>Life expectancy of at least 6 months</td>
</tr>
<tr>
<td>Fluent in English</td>
</tr>
<tr>
<td>Have home internet access</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Exclusion criteria</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical evidence of cognitive or psychological impairment</td>
</tr>
<tr>
<td>Prisoners and pregnant women</td>
</tr>
<tr>
<td>Currently participating in other psychosocial studies</td>
</tr>
</tbody>
</table>

Study recruitment included physician referrals, advertisements and flyers, and a patient health care portal (SupportScreen [45]) from City of Hope. Participants were enrolled between October 2017 and September 2019. All study participants were screened for complete eligibility criteria and provided written informed consent before study participation. Consented participants were randomized to either the PatientCareAnywhere intervention or usual care control condition using a computer-based random assignment program using a 1 to 1 ratio. Due to the nature of the study, it was not possible to blind participants’ study conditions. Participants in both the intervention and control groups participated in their respective study arm for a 3-month period and completed a baseline assessment at the time of enrollment (T1), which included a sociodemographic questionnaire and 2 biopsychosocial questionnaires assessing HRQOL, as measured by the Functional Assessment of Cancer Therapy-General (FACT-G) [51], and cancer-related self-efficacy, as measured by the Self-Efficacy for Managing Chronic Disease (SEMCD) [52]. Follow-up assessments (FACT-G and SEMCD) were completed monthly until the end of participation, resulting in 3 additional time points: 4 weeks from baseline (T2), 8 weeks from baseline (T3), and 12 weeks from baseline (T4). Table 2 outlines the procedures conducted at different time points of the RCT.
### Table 2. Procedures conducted at different phases of the pilot randomized controlled trial.

<table>
<thead>
<tr>
<th>Time</th>
<th>Patient tasks</th>
<th>Clinical team tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment</td>
<td>• All: complete sociodemographic questionnaire at clinic</td>
<td>• All: enter clinic data into trial management system</td>
</tr>
<tr>
<td></td>
<td>• All: complete baseline FACT-G&lt;sup&gt;a&lt;/sup&gt; and SEMCD&lt;sup&gt;b&lt;/sup&gt; (T1&lt;sup&gt;c&lt;/sup&gt;)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• All: set up an account for and receive an orientation to the PatientCareAnywhere system</td>
<td></td>
</tr>
<tr>
<td>Once a week</td>
<td>• Intervention: use the PatientCareAnywhere system, including the symptom reporting feature</td>
<td>• None</td>
</tr>
<tr>
<td>Every month</td>
<td>• All: complete FACT-G and SEMC after 1 month (T2&lt;sup&gt;d&lt;/sup&gt;) and 2 months (T3&lt;sup&gt;e&lt;/sup&gt;) of participation</td>
<td>• All: enter survey data collected on paper into the system</td>
</tr>
<tr>
<td>Before every visit</td>
<td>• None</td>
<td>• Intervention: ensure symptom assessment report from PatientCareAnywhere is either printed or available on computer</td>
</tr>
<tr>
<td>During every visit</td>
<td>• Intervention: review symptom assessment report with the provider</td>
<td>• Intervention: review symptom assessment report with the patient</td>
</tr>
<tr>
<td>Conclusion of participation</td>
<td>• All: complete FACT-G and SEMCD after 3 months of participation (end of the study; T4&lt;sup&gt;f&lt;/sup&gt;)</td>
<td>• All: enter all paper-based data into trial management system</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Intervention: compile metrics of PatientCareAnywhere system use (eg, frequency of use and time spent)</td>
</tr>
</tbody>
</table>

<sup>a</sup>FACT-G: Functional Assessment of Cancer Therapy-General.

<sup>b</sup>SEMCD: Self-Efficacy for Managing Chronic Disease.

<sup>c</sup>At the time of enrollment (baseline).

<sup>d</sup>4 weeks after baseline (first midpoint of the study).

<sup>e</sup>8 weeks after baseline (second midpoint of the study).

<sup>f</sup>12 weeks after baseline (end of participation).

### Study Conditions

#### Intervention Condition
Patients in the intervention group were encouraged to use PatientCareAnywhere at least weekly to not only report symptoms when necessary but also use other features of the site, such as the education content. Reminder emails were sent to patients to encourage the use of the system after 1 week of inactivity. Patient-reported symptoms of moderate or worse severities triggered email alerts to the study coordinators for triage, who then contacted appropriate providers or supportive care staff to address the patient’s concerns.

#### Control Condition
Patients in the control group received usual care, including a 1-time use of SupportScreen for symptom checking at the clinic during initial treatment consultation after a cancer diagnosis. The use of SupportScreen could also trigger the delivery of consultation, print patient education materials, and specialist referrals.

### Measures

#### Sociodemographic and Cancer-Specific Characteristics
At baseline, before the intervention, patients self-reported sociodemographic information (eg, age, race, ethnicity, education, and income) and clinical information (eg, cancer diagnosis and stage of cancer), which were confirmed via medical record review.

#### Health-Related Quality of Life
The FACT-G is a 27-item self-report questionnaire designed to measure 4 domains of HRQOL in patients with cancer, including emotional, functional, physical, and social well-being [51]. Patients rate the degree to which the items applied to them over the past 7 days using a 5-point response scale ranging from 1=“not at all” to 5=“very much.” Total FACT-G scores range from 0 to 108, with a higher score indicating better quality of life.

#### Patient Self-Efficacy
Patient self-efficacy is an essential component of the treatment and management of illnesses, including cancer. The 6-item SEMCD scale measures patients’ confidence in their ability to manage fatigue, physical discomfort or pain, emotional distress, and other symptoms or health problems; to carry out different tasks or activities to reduce the need to see a physician; and to do things other than taking medication to reduce illness effects [52,53]. Items are rated on a 10-point scale ranging from 1=“not at all confident” to 10=“totally confident,” and scores are averaged across items. The final score (mean of the 6 items) ranges from 0 to 10, with higher scores indicating greater self-efficacy.
Intervention Use

PatientCareAnywhere tracked the frequency with which participants accessed the intervention over the 3-month study period. The system also recorded participants’ responses to multiple symptom assessments and the time (minutes) it took to complete each assessment.

Sample Size

The primary goal of the pilot RCT was to compare the FACT-G change across time in the intervention group with the FACT-G change across time in the control group. The sample size calculation was based on prior research that established the minimally important difference for the total FACT-G ranges from 4 to 7 points [54-56]. Specifically, a sample size of 72 participants (36 participants per group) would achieve >80% power to detect a difference in mean changes of 7 (with SD of 12 at both time points and a correlation between measurement pairs of 0.65). The significance level is .05 using a 2-sided, 2-sample t test.

Data Analytic Plan

Descriptive statistics (eg, means, frequencies, and percentages) were used to characterize the sociodemographic and disease characteristics of the RCT participants. Demographic differences between intervention and control groups were evaluated using t test for continuous variables and Fisher exact test for categorical variables. Regarding FACT-G and SEMCD scores, independent sample t test was used to compare mean differences (ie, mean difference between pre- and postintervention scores) between the 2 groups at T4. All statistical analyses were conducted using SPSS.

Hypotheses

We hypothesized that at postintervention, patients randomized to the PatientCareAnywhere intervention would report better HRQOL outcomes, as measured by the FACT-G (primary hypothesis), and self-efficacy, as measured by the SEMCD (secondary hypothesis), compared with patients randomized to the usual care control condition.

Results

Phase 1: Usability Testing

Participant Characteristics

A total of 11 participants (patients: n=4 and caregivers: n=7) participated in usability testing with a prototype of the PatientCareAnywhere system. This sample size was justified based on previous usability research demonstrating that 5 to 7 participants is sufficient to reveal about 80% of the usability issues [57]. Table 3 presents the sociodemographic characteristics of the usability testing sample. Patients were mostly non-Hispanic (7/11, 64%) and White (10/11, 91%), with an average age of 50 (SD 6.8) years. The average age of caregivers was 44 (SD 20) years.

Table 3. Sample characteristics of usability testing participants (N=11).

<table>
<thead>
<tr>
<th></th>
<th>Patients (n=4)</th>
<th>Caregivers (n=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (y), mean (SD; range)</strong></td>
<td>50.3 (6.8; 44-59)</td>
<td>43.7 (20.4; 23-73)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>0 (0)</td>
<td>4 (57.14)</td>
</tr>
<tr>
<td>Woman</td>
<td>4 (100)</td>
<td>3 (42.86)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1 (25)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>White</td>
<td>3 (75)</td>
<td>7 (100)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>2 (50)</td>
<td>1 (14.29)</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>2 (50)</td>
<td>5 (71.43)</td>
</tr>
<tr>
<td>Unknown</td>
<td>0 (0)</td>
<td>1 (14.29)</td>
</tr>
</tbody>
</table>

Usability Outcomes

Qualitative Results

Individual interviews with patients with cancer and caregivers were conducted to evaluate the usability and usefulness of the PatientCareAnywhere system. The interviews consisted of asking the participants to complete a list of tasks that addressed each of the features of the PatientCareAnywhere system (eg, where to find certain information on the page or how to complete a symptom report) and recording the time it took for each task to be completed as well as identifying any tasks that were difficult to complete. Multimedia Appendix 3 provides the list of tasks that were asked of participants. In addition to the task-completion activity, we solicited feedback from the participants on the site functions, features, and design as well as their ideas for improvement. All participants were able to complete the tasks within 5 seconds of being asked, and no participant experienced confusion about navigating the site and completing specific activities.

Participants also had high levels of satisfaction with the PatientCareAnywhere design, features, and functionality of the system. Specifically, patients enjoyed the ability to connect with friends, family, community organizations, other patients with cancer and survivors of cancer, and their care team. They felt
that the PatientCareAnywhere layout made sense to them as users, and they did not find any parts of the pages to be confusing. The patients uniformly liked the layout of the biopsychosocial screening tool and preferred the idea that they only had to give responses to the topics that they were concerned about on the tool. They also felt that the personalized recommendations would be a great asset to them during their cancer journey and were especially happy about being able to report their symptoms at any time. In addition, patients really liked the “one-stop-shop” idea of PatientCareAnywhere—the ability to keep track of their appointments and medical information in the same place as connecting with friends and family and finding local events and support groups. All patients felt that the wireframes were well thought out, and each one asked when the system would be available for use at City of Hope.

The caregivers also highly praised the PatientCareAnywhere wireframes. All caregiver participants felt that the wireframes were laid out in a logical manner. They particularly liked having access to their loved one’s medical records and appointments (given only with the caregiver permission level), and they felt that this system would make caregiving a much easier experience. Other features that the caregivers highlighted would make a difference for them were the ability to complete a symptom report for their loved one (patient) and the “help request” feature that would allow caregivers (and patients) to send requests for help (eg, assistance with transportation) to their PatientCareAnywhere friends. The PatientCareAnywhere friends can respond to the email request if they can help and this affirmation is noted by the PatientCareAnywhere system. Overall, patients and caregivers did not have any trouble identifying how to complete the biopsychosocial screening tool and where to find recommendations, medical information, educational materials, and local events on the PatientCareAnywhere wireframes.

Quantitative Results

The average SUS total score was 84.09 (SD 10.02; range 75.00-100.00), which was well above the predetermined 70-point threshold reflecting “excellent” usability (Table 4). Regarding the Usefulness Questionnaire, participants agreed or strongly agreed with all 35 statements (refer to Tables 5 and 6, which include average ratings for each statement). Specifically, patients and caregivers rated the usefulness of PatientCareAnywhere site features from 4.00 to 5.00 (Table 5) or 3.86 to 4.86 (Table 6) and, respectively (4=“agree” and 5=“strongly agree”).

Table 4. Results from the System Usability Scale questionnaire (N=11)a.

<table>
<thead>
<tr>
<th>Item</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I think that I would like to use this system frequently.</td>
<td>4.18 (0.60)</td>
</tr>
<tr>
<td>2. I found the system unnecessarily complex.</td>
<td>1.64 (0.82)</td>
</tr>
<tr>
<td>3. I thought the system was easy to use.</td>
<td>4.36 (0.48)</td>
</tr>
<tr>
<td>4. I think that I would need the support of a technical person to be able to use this system.</td>
<td>1.36 (0.70)</td>
</tr>
<tr>
<td>5. I found the various functions in this system were well integrated.</td>
<td>4.36 (0.48)</td>
</tr>
<tr>
<td>6. I thought there was too much inconsistency in this system.</td>
<td>1.55 (0.70)</td>
</tr>
<tr>
<td>7. I would imagine that most people would learn to use this system very quickly.</td>
<td>4.27 (0.42)</td>
</tr>
<tr>
<td>8. I found the system very cumbersome to use.</td>
<td>1.73 (1.03)</td>
</tr>
<tr>
<td>9. I felt very confident using the system.</td>
<td>4.36 (0.67)</td>
</tr>
<tr>
<td>10. I needed to learn a lot of things before I could get going with this system.</td>
<td>1.64 (0.67)</td>
</tr>
</tbody>
</table>

aTotal usability score is a sum of individual items multiplied by 2.5 to convert original scores of 0 to 40 to 0 to 100. Possible item responses range from 1 (strongly disagree) to 5 (strongly agree).
Table 5. Results from the Patient’s Usefulness Questionnaire (N=4)\textsuperscript{a}.

<table>
<thead>
<tr>
<th>Item number and item</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In managing my cancer care, it is useful to...</td>
<td></td>
</tr>
<tr>
<td>1. Connect with friends, family, and doctors/nurses through private messaging and</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>wall posting features.</td>
<td></td>
</tr>
<tr>
<td>2. Receive education recommendations that are tailored to my medical situation and/or</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>personal needs.</td>
<td></td>
</tr>
<tr>
<td>3. View support group recommendations that are tailored to my needs.</td>
<td>4.25 (0.50)</td>
</tr>
<tr>
<td>4. View recommendations for local classes and events that are tailored to my needs.</td>
<td>4.25 (0.50)</td>
</tr>
<tr>
<td>5. Be able to create help requests that are sent out to caregivers and/or friends.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>6. Report symptoms via the symptom reporting tool.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>7. Be able to track my symptoms over time via the symptom reporting tool.</td>
<td>4.25 (0.50)</td>
</tr>
<tr>
<td>8. View the educational articles that were recommended to me based off the reported</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>symptoms.</td>
<td></td>
</tr>
<tr>
<td>9. Have access to my medication and supplement list.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>10. Have access to my laboratories and tests results.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>11. Have access to my other medical records.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>12. Be able to add additional medical information or upload other medical documents.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>13. View the care team members that have received referrals regarding my personal or</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>medical needs.</td>
<td></td>
</tr>
<tr>
<td>14. See the events that are scheduled at the City of Hope.</td>
<td>4.00 (0.82)</td>
</tr>
<tr>
<td>15. See the events for which I am registered.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>16. Add my own events to my calendar.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>17. View the medical appointments that are scheduled for me.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>18. View the help requests that have been sent out.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>19. See which classes, events and support groups are available at the City of Hope.</td>
<td>4.25 (0.50)</td>
</tr>
<tr>
<td>20. See which classes, events and support groups are available in my local area.</td>
<td>4.25 (0.50)</td>
</tr>
<tr>
<td>21. Be able to register for a class, event or support group.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>22. Read a description of the class/event/support group and the event leader’s</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>contact info.</td>
<td></td>
</tr>
<tr>
<td>23. Be able to read the educational content/articles that have been recommended to</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>me.</td>
<td></td>
</tr>
<tr>
<td>24. Be able to save articles that I want to reference later into a “Favorites” area.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>25. Be able to browse educational materials by category.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>26. Be able to request additional information about a topic.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>27. Be able to share an educational article with the PCA\textsuperscript{b}</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>administrators so that they could add it to PCA.</td>
<td></td>
</tr>
<tr>
<td>In general, I feel...</td>
<td></td>
</tr>
<tr>
<td>1. Comfortable using PCA on my own.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>2. That PCA is an easy site to navigate.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>3. That the overall look-and-feel of PCA is appealing.</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>4. That the overall organization of PCA is logical.</td>
<td>5.00 (0.00)</td>
</tr>
<tr>
<td>5. That for noncritical medical situations, I would rather get information and nurse</td>
<td>4.50 (0.58)</td>
</tr>
<tr>
<td>help via PCA instead of having an in-person doctor’s appointment.</td>
<td></td>
</tr>
<tr>
<td>6. That I would recommend PCA to other caregivers.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>7. That I would recommend PCA to other patients.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>8. That cancer centers should use PCA as part of their standard care practices.</td>
<td>4.75 (0.50)</td>
</tr>
<tr>
<td>Total score</td>
<td>4.66 (0.28)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}The highest score is 5.00 (strongly agree) and the lowest score is 1.00 (strongly disagree).

\textsuperscript{b}PCA: PatientCareAnywhere.
Table 6. Results from the Caregiver’s Usefulness Questionnaire (N=7).\(^a\)

<table>
<thead>
<tr>
<th>In caring for my family member/friend who has cancer, it is useful to...</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Connect with her/him and others through private messaging and wall posting features.</td>
<td>3.86 (0.90)</td>
</tr>
<tr>
<td>2. Receive education recommendations that are tailored to my friend/family member’s medical situation.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>3. View support group recommendations that are tailored to my friend/family member’s needs.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>4. View recommendations for local classes and events that are tailored to my friend/family member’s needs.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>5. Be able to create help requests that are sent out to other caregivers and/or friends.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>6. Report symptoms via the symptom reporting tool on behalf of my family member/friend.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>7. Be able to track her/his symptoms over time via the symptom reporting tool.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>8. View the educational articles that were recommended for her/him based off the reported symptoms.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>9. Have access to her/his medication and supplement list.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>10. Have access to her/his laboratories and tests results.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>11. Have access to her/his other medical records.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>12. Be able to add additional medical information or upload other medical documents.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>13. View the care team members that have received referrals for my family member/friend.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>14. See the events that are scheduled at the city of hope.</td>
<td>4.43 (0.79)</td>
</tr>
<tr>
<td>15. See the events for which he/she is registered.</td>
<td>4.43 (0.53)</td>
</tr>
<tr>
<td>16. Add my own events to my calendar.</td>
<td>4.29 (0.95)</td>
</tr>
<tr>
<td>17. Add events for my family member/friend.</td>
<td>4.29 (0.95)</td>
</tr>
<tr>
<td>18. View the medical appointments that are scheduled for my family member/friend.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>19. View the help requests that have been sent out.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>20. See which classes, events and support groups are available at the city of hope.</td>
<td>4.43 (0.98)</td>
</tr>
<tr>
<td>21. See which classes, events and support groups are available in my local area.</td>
<td>4.43 (0.79)</td>
</tr>
<tr>
<td>22. View the classes, events and support groups that are recommended for my family member/friend.</td>
<td>4.43 (0.79)</td>
</tr>
<tr>
<td>23. Be able to register my family member/friend for a class, event or support group on their behalf.</td>
<td>4.29 (0.76)</td>
</tr>
<tr>
<td>24. Read a description of the class/event/support group and the event leader’s contact info.</td>
<td>4.43 (0.79)</td>
</tr>
<tr>
<td>25. Be able to read the educational content/articles that have been recommended to my family member/friend.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>26. Be able to save articles that i want to reference later into a “favorites” area.</td>
<td>4.86 (0.38)</td>
</tr>
<tr>
<td>27. Be able to browse educational materials by category.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>28. Be able to request additional information about a topic.</td>
<td>4.71 (0.49)</td>
</tr>
<tr>
<td>29. Be able to share an educational article with the PCA(^b) administrators so that they could add it to PCA.</td>
<td>4.43 (0.53)</td>
</tr>
</tbody>
</table>

For me personally as a caregiver, it is useful to...

1. Have education recommendations that are tailored to my role as caregiver | 4.86 (0.38) |
2. Have support group recommendations that are tailored to my role as caregiver. | 4.29 (0.76) |
3. Have recommendations for local classes and events that are tailored to my role as caregiver. | 4.43 (0.53) |

In general, I feel...

1. Comfortable using PCA on my own. | 4.57 (0.79) |
2. That PCA is an easy site to navigate. | 4.43 (0.79) |
3. That the overall look-and-feel of PCA is appealing. | 4.43 (0.79) |
4. That the overall organization of PCA is logical. | 4.57 (0.53) |
5. That for noncritical medical situations, I would rather get information and nurse help via PCA instead of having an in-person doctor’s appointment. | 4.29 (0.76) |
6. That I would recommend PCA to other caregivers. | 4.71 (0.76) |
Phase 2: Pilot RCT

Participant Characteristics

A total of 72 patients with cancer were enrolled and individually randomized (1:1) to the PatientCareAnywhere intervention (n=36, 50%) or usual care control condition (n=36, 50%) for 3 months. The following analysis was limited to 59 participants who completed at least 2 of the questionnaires (FACT-G and SEMCD): 28 (47%) patients in the intervention group and 31 (53%) patients in the control group. Of note, there were no significant differences in demographic characteristics between the included (59/72, 82%) and excluded (13/72, 18%) participants (Multimedia Appendix 4). Table 7 summarizes the pilot RCT participants’ sociodemographic and cancer-related characteristics, with no significant between-group differences. Overall, the RCT participants had a mean age of 53.85 (SD 12.37) years and were predominantly women (49/59, 83%), White (41/59, 69%), and non-Hispanic or Latino (41/59, 69%). Most participants were married (43/59, 73%), had at least a college degree (33/59, 56%), and earned >US $100,000 (29/59, 49%). The most common diagnosis was breast cancer (43/59, 73%) and nonmetastatic (stages 0-III; 33/59, 56%).

<table>
<thead>
<tr>
<th>Score, mean (SD)</th>
<th>7. That I would recommend PCA to other patients.</th>
<th>4.71 (0.76)</th>
<th>8. That cancer centers should use PCA as part of their standard care practices.</th>
<th>4.57 (0.79)</th>
<th>Total scores</th>
<th>4.59 (0.23)</th>
</tr>
</thead>
</table>

The highest score is 5.00 (strongly agree) and the lowest score is 1.00 (strongly disagree).

PCA: PatientCareAnywhere.
Table 7. Sample characteristics of the pilot randomized controlled trial participants (N=59).

<table>
<thead>
<tr>
<th></th>
<th>Intervention (n=28)</th>
<th>Control (n=31)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (y), mean (SD; range)</strong></td>
<td>54.82 (12.32; 34-77)</td>
<td>52.97 (12.37; 30-79)</td>
<td>.57</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>4 (14.29)</td>
<td>6 (19.35)</td>
<td>.73</td>
</tr>
<tr>
<td>Woman</td>
<td>24 (85.71)</td>
<td>25 (80.65)</td>
<td></td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
<td></td>
<td>.55</td>
</tr>
<tr>
<td>Asian American</td>
<td>6 (21.43)</td>
<td>4 (12.90)</td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>2 (7.14)</td>
<td>5 (16.13)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>20 (71.43)</td>
<td>21 (67.74)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>1 (3.23)</td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
<td></td>
<td>.24</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>9 (32.14)</td>
<td>6 (19.35)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>17 (60.71)</td>
<td>24 (77.42)</td>
<td></td>
</tr>
<tr>
<td>Unknown or not reported</td>
<td>2 (7.14)</td>
<td>1 (3.23)</td>
<td></td>
</tr>
<tr>
<td><strong>Marital status, n (%)</strong></td>
<td></td>
<td></td>
<td>.07</td>
</tr>
<tr>
<td>Single</td>
<td>2 (7.14)</td>
<td>6 (19.35)</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>19 (67.86)</td>
<td>24 (77.42)</td>
<td></td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>5 (17.86)</td>
<td>1 (3.23)</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>2 (7.14)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Education, n (%)</strong></td>
<td></td>
<td></td>
<td>.23</td>
</tr>
<tr>
<td>Less than high school</td>
<td>2 (7.14)</td>
<td>1 (3.23)</td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>5 (17.86)</td>
<td>7 (22.58)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>13 (46.43)</td>
<td>20 (64.52)</td>
<td></td>
</tr>
<tr>
<td>Graduate school</td>
<td>8 (28.57)</td>
<td>3 (9.68)</td>
<td></td>
</tr>
<tr>
<td><strong>Household income (US $), n (%)</strong></td>
<td></td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>&lt;20,000</td>
<td>4 (14.29)</td>
<td>4 (12.90)</td>
<td></td>
</tr>
<tr>
<td>20,000-29,999</td>
<td>2 (7.14)</td>
<td>1 (3.23)</td>
<td></td>
</tr>
<tr>
<td>30,000-49,999</td>
<td>2 (7.14)</td>
<td>4 (12.90)</td>
<td></td>
</tr>
<tr>
<td>50,000-69,999</td>
<td>3 (10.71)</td>
<td>5 (16.13)</td>
<td></td>
</tr>
<tr>
<td>70,000-99,999</td>
<td>3 (10.71)</td>
<td>2 (6.45)</td>
<td></td>
</tr>
<tr>
<td>&gt;100,000</td>
<td>14 (50)</td>
<td>15 (48.39)</td>
<td></td>
</tr>
<tr>
<td><strong>Cancer, n (%)</strong></td>
<td></td>
<td></td>
<td>.57</td>
</tr>
<tr>
<td>Breast</td>
<td>20 (71.43)</td>
<td>23 (74.19)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>5 (17.86)</td>
<td>2 (6.45)</td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>2 (7.14)</td>
<td>3 (9.68)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (3.57)</td>
<td>3 (9.68)</td>
<td></td>
</tr>
<tr>
<td><strong>Disease stage, n (%)</strong></td>
<td></td>
<td></td>
<td>.86</td>
</tr>
<tr>
<td>Stage 0</td>
<td>2 (7.14)</td>
<td>3 (9.68)</td>
<td></td>
</tr>
<tr>
<td>Stage 1</td>
<td>6 (21.43)</td>
<td>5 (16.13)</td>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
<td>4 (14.29)</td>
<td>7 (22.58)</td>
<td></td>
</tr>
<tr>
<td>Stage 3</td>
<td>2 (7.14)</td>
<td>4 (12.90)</td>
<td></td>
</tr>
<tr>
<td>Stage 4</td>
<td>8 (28.57)</td>
<td>7 (22.58)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>6 (21.43)</td>
<td>5 (16.13)</td>
<td></td>
</tr>
</tbody>
</table>
Study Outcomes

Health-Related Quality of Life
The mean difference in FACT-G scores between the preintervention (T1; baseline) and postintervention (T4; 12-week postbaseline) assessments of each patient (Table 8) for the intervention group was 6.08 (SD 15.26), indicating an improvement in HRQOL among patients who received PatientCareAnywhere. For the control group, the mean difference in FACT-G scores between the preintervention (T1) and postintervention (T4) assessments was –2.95 (SD 10.63), indicating a worsening of HRQOL among patients who received usual care. The between-group difference was statistically significant (P=.01), with a medium effect size (Cohen d=0.70).

Table 8. Results from the pilot randomized controlled trial.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean difference± (SD)</th>
<th>P value\textsuperscript{b}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HRQOL\textsuperscript{c} (FACT-G\textsuperscript{d})</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n=31)</td>
<td>–2.95 (10.63)</td>
<td>.01</td>
</tr>
<tr>
<td>Intervention (n=26)</td>
<td>6.08 (15.26)</td>
<td>.01</td>
</tr>
<tr>
<td><strong>Self-efficacy (SEMCD\textsuperscript{e})</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n=31)</td>
<td>–0.84 (11.20)</td>
<td>.09</td>
</tr>
<tr>
<td>Intervention (n=27)</td>
<td>4.22 (10.91)</td>
<td>.09</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Mean difference between preintervention (T1; baseline) and postintervention (T4; 12 weeks from baseline) scores.
\textsuperscript{b} Significant P values (P<.05) are italicized.
\textsuperscript{c} HRQOL: health-related quality of life.
\textsuperscript{d} FACT-G: Functional Assessment of Cancer Therapy-General.
\textsuperscript{e} SEMCD: Self-Efficacy for Managing Chronic Disease.

Patient Self-Efficacy
Similarly, the mean difference in SEMCD scores between the preintervention (T1; baseline) and postintervention (T4; 12-week postbaseline) assessments (Table 8) for the intervention group was 4.22 (SD 10.91) and for the control group was –0.84 (SD 11.20). However, the between-group difference was not statistically significant (P=.09), with a small-to-medium effect size (Cohen d=0.46).

Intervention Use
Overall, 61% (17/28) of the patients in the intervention group were classified as “Frequent Users,” defined as having accessed the PatientCareAnywhere site at least 5 times during the study. Among the frequent users, the mean difference between the first and last FACT-G scores was 7.12 (SD 15.4), which was statistically significantly higher than that of the control group (P=.007), with a large effect size (Cohen d=0.80). The mean difference between the first and last SEMCD scores (mean 5.47, SD 6.43) was also statistically significantly better than that of the control group (P=.03), with a medium effect size (Cohen d=0.71). In comparison, among the infrequent users (n=11), the mean difference between the first and last scores on the FACT-G (mean 4.02, SD 17.39; P=.10) and SEMCD (mean 0.73, SD 15.37; P=.68) did not significantly differ from the control group.

Symptom Reporting
Finally, a total of 140 symptom reports were recorded. On average, each symptom reporting session included 4.4 symptoms and lasted for 3.4 minutes.

Discussion
The primary aim of this study was to evaluate the usability and preliminary efficacy of PatientCareAnywhere, a patient-centered adaptive supportive care system, to improve patient-reported outcomes for patients newly diagnosed with cancer.

Principal Findings
Both qualitative and quantitative usability testing results were notably positive and support the usability of PatientCareAnywhere. Overall, patients with cancer and caregivers were highly satisfied with the purpose and functions of the intervention, found the content relevant and useful, and expressed strong support for the biopsychosocial screening tool and personalized recommendations. On the basis of participant feedback, several changes were made to the design features: a “help request” function was added, caregivers were given greater access to the patient’s medical information, and symptom reporting was added to the caregiver portal, allowing caregivers to report symptoms on the patient’s behalf. In addition, the quantitative feedback demonstrated a high usability level for PatientCareAnywhere, with an average SUS score of 84.09 (SD 10.02), indicating above-average acceptable usability [49,50]. The scores on the Usefulness Questionnaire also reflected the positive experience that users had with the system and underlined the beliefs of participants that the features of PatientCareAnywhere were acceptable and useful during the cancer care journey.

Results from the pilot RCT demonstrate the preliminary efficacy of the PatientCareAnywhere intervention. Compared with usual care, patients with cancer who received PatientCareAnywhere showed statistically significant improvements in HRQOL from pre- to postintervention scores. While self-efficacy scores also increased in the intervention group, the difference was not statistically significant compared with the control group. When evaluating intervention use, “frequent users” (ie, patients who accessed the intervention at least 5 times during the study)
reported greater improvements in both HRQOL and self-efficacy outcomes (medium to large effect sizes) compared with the control group. These results confirmed our hypotheses that routine use of PatientCareAnywhere could result in improved HRQOL outcomes and greater patient self-efficacy among patients newly diagnosed with cancer, and that these effects were more prominent with greater intervention use. Furthermore, patients on average reported about 4 symptoms and completed the symptom assessment in <4 minutes. Notably, this is drastically shorter than SupportScreen, which takes approximately 15 to 20 minutes to complete [45]. This observation indicates that PatientCareAnywhere is also an efficient symptom reporting tool.

There is growing evidence of technology-assisted assessments and interventions enhancing the delivery of patient-centered care through improved symptom monitoring, communication between patients and providers, tailored resources, and patient empowerment and engagement across the continuum of care [42]. Our findings are in line with previous studies that have demonstrated the effectiveness of technology-based interventions in cancer care [42]. Recently, a comprehensive scoping review was conducted on 134 literature reviews of digital health and telehealth interventions across the cancer continuum, in which a majority focused on patients with cancer (n=128) in the active treatment (n=48) and survivorship (n=29) phases using eHealth programs, synchronous telehealth, mobile apps, asynchronous messaging (eg, email), and SMS text messaging [58]. A total of 29 reviews included a meta-analysis, with results signifying positive effects of digital health and telehealth in cancer care on quality of life, psychological outcomes (eg, anxiety and depression), and cancer screenings [58]. Of note, the benefits of digital supportive cancer care interventions have been demonstrated independent of demographic and disease factors [44,59]. The lack of a positive effect on self-efficacy warrants further evaluation. Similar to our findings, an RCT evaluating an internet-based interactive health communication application that allows patients with cancer to monitor their symptoms and provides tailored self-management support reported no significant between-group differences in depression, HRQOL, self-efficacy, and social support, although self-efficacy and HRQOL outcomes significantly worsened over time in the control group [60]. Conversely, an earlier review and meta-analysis of eHealth-based self-management interventions demonstrated a statistically significant effect on self-efficacy but noted that the effect size was small (0.4) [61]. A larger sample size may be needed to observe meaningful changes in self-efficacy.

**Strengths and Limitations of PatientCareAnywhere**

PatientCareAnywhere was developed using a user-centered design approach to ensure the needs and preferences of patients newly diagnosed with cancer were addressed. Applying user-centered design principles to the overall development of PatientCareAnywhere resulted in a user-friendly, functional, useful, and acceptable supportive care intervention. In addition, the web-based delivery and responsive-design technologies allow patients to access the intervention at anytime and anywhere, including outside of clinic and at home, and use PatientCareAnywhere on multiple devices (eg, desktops, smartphones, and tablets), thereby providing more flexible intervention delivery and reducing common practical barriers to care (eg, transportation issues and scheduling conflicts). Furthermore, by remotely and routinely monitoring patients’ biopsychosocial symptoms and distress, PatientCareAnywhere provides supportive cancer care tailored to their needs. Compared with other distress management systems, additional advantages of PatientCareAnywhere include (1) PatientCareAnywhere has access to numerous validated questionnaires, allowing an institution to pick and choose the ones that are most suitable for their patients and providers; (2) PatientCareAnywhere allows the invocation of another questionnaire for follow-up questions based on the results from a top-level screening tool; (3) PatientCareAnywhere provides a communication platform to allow caregivers, family members, and friends to interact directly with the patient via the system; (4) PatientCareAnywhere allows community organizations to post events and respond to patient requests for help; (5) PatientCareAnywhere delivers tailored and responsive patient education contents that evolve with the patient based on their current needs and concerns; and (6) PatientCareAnywhere is backed by City of Hope’s comprehensive supportive care training program.

This study has some limitations. First, the objectives of this study were to establish the usability and preliminary effects of PatientCareAnywhere rather than investigate intervention efficacy. However, results from this pilot study will inform the next phase of research to conduct a full-scale RCT evaluating the efficacy of PatientCareAnywhere to improve HRQOL and reduce symptom burden compared with usual care. Second, the study was limited to patients diagnosed with breast, lung, or prostate cancer. We initially focused on these 3 common cancers to develop cancer-specific educational materials, with plans to expand to all cancer types (eg, gastrointestinal, gynecologic, head and neck, and hematologic) and treatment options following successful initial pilot testing results. Further study is warranted to adapt PatientCareAnywhere to other types of cancer and examine the extent to which our initial findings are generalizable to patients with different cancer diagnoses. Third, while sufficient for the purposes of our study, sample sizes for usability testing (N=11) and pilot RCT (N=78) were relatively small, limiting power and study results. It is possible that with a larger sample, stronger intervention effects may emerge. Fourth, study participants were limited to patients receiving care at City of Hope and may not be representative of the general cancer population in the United States. Furthermore, the intervention was only available in English and required patients to have internet access to participate in the study, which also may limit the generalizability of our findings. Future studies should investigate the efficacy of PatientCareAnywhere with a more diverse and larger sample of patients with cancer over a longer study period and explore the optimal intervention use to improve patient outcomes.

**Conclusions**

Engaging key stakeholders throughout the design and development process helped ensure PatientCareAnywhere was a patient-centered, user-friendly, efficient, and effective supportive care system. Overall, the results from initial pilot
testing demonstrate the usability and preliminary efficacy of PatientCareAnywhere to improve HRQOL outcomes among patients newly diagnosed with cancer.

Acknowledgments
Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health (R44CA192588). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability
The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions
Conceptualization of this paper was done by SHB, KC, ML, DY, WD, and NH. SHB contributed to data analysis and interpretation, writing, and original draft preparation. KC, ML, DY, and NH participated in obtaining funding, methodology, project administration, and supervision of the study. AC was involved in project implementation, including participant recruitment, study coordination, and collection and assembly of data. MR conducted the formal analysis. KC, MS, ML, MR, WD, and NH contributed to the interpretation of data and writing, reviewing, and editing of the original manuscript. All authors read and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Main components of PatientCareAnywhere.
[DOCX File, 18 KB - cancer_v10i1e49703_app1.docx]

Multimedia Appendix 2
Screenshots of PatientCareAnywhere.
[DOCX File, 2954 KB - cancer_v10i1e49703_app2.docx]

Multimedia Appendix 3
PatientCareAnywhere usability tasks (phase 1).
[DOCX File, 18 KB - cancer_v10i1e49703_app3.docx]

Multimedia Appendix 4
Sample characteristics of the included versus excluded pilot randomized controlled trial participants (phase 2).
[DOCX File, 20 KB - cancer_v10i1e49703_app4.docx]

Multimedia Appendix 5
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2119 KB - cancer_v10i1e49703_app5.pdf]

References


https://cancer.jmir.org/2024/1/e49703


Abbreviations

- EHR: electronic health record
- FACT-G: Functional Assessment of Cancer Therapy-General
- HRQOL: health-related quality of life
- NCCN: National Comprehensive Cancer Network
- RCT: randomized controlled trial
- SEMCD: Self-Efficacy for Managing Chronic Disease
- SUS: System Usability Scale

Edited by T de Azevedo Cardoso; submitted 06.06.23; peer-reviewed by R Granger, R Vanderpool, P Innominato, R Vanderpool; comments to author 19.07.23; revised version received 22.12.23; accepted 21.03.24; published 10.07.24.

Please cite as:
Baik SH, Clark K, Sanchez M, Loscalzo M, Celis A, Razavi M, Yang D, Dale W, Haas N

Usability and Preliminary Efficacy of an Adaptive Supportive Care System for Patients With Cancer: Pilot Randomized Controlled Trial
JMIR Cancer 2024;10:e49703
URL: https://cancer.jmir.org/2024/1/e49703
doi:10.2196/49703
PMID:
Toxic Relationships Described by People With Breast Cancer on Reddit: Topic Modeling Study

Cara Anne Davidson¹, MSc; Richard Booth², PhD; Kimberley Teresa Jackson², PhD; Tara Mantler³, PhD

¹Department of Health and Rehabilitation Sciences, Faculty of Health Sciences, Western University, London, ON, Canada
²Arthur Labatt Family School of Nursing, Faculty of Health Sciences, Western University, London, ON, Canada
³School of Health Studies, Faculty of Health Sciences, Western University, London, ON, Canada

Corresponding Author:
Cara Anne Davidson, MSc
Department of Health and Rehabilitation Sciences
Faculty of Health Sciences
Western University
1151 Richmond St
London, ON, N6A 3K7
Canada
Phone: 1 519 661 2111 ext 85541
Email: cda davidson@uw o.ca

Abstract

Background: Social support is essential to promoting optimal health outcomes for women with breast cancer. However, an estimated 12% of women with breast cancer simultaneously experience intimate partner violence (IPV; physical, psychological, or sexual abuse by an intimate partner). Women who experience IPV during breast cancer may lack traditional social support, and thus seek out alternative sources of support. Online community forums, such as Reddit, can provide accessible social connections within breast cancer–specific communities. However, it is largely unknown how women with breast cancer use Reddit to describe and seek support for experiences of IPV.

Objective: This study aims to explore how patients with breast cancer describe toxic relationships with their partners and immediate family members on Reddit.

Methods: This exploratory, cross-sectional, topic-modeling study analyzed textual data from 96 users in the r/breastcancer subreddit in February 2023. The meaning extraction method, inclusive of principal component analysis, was used to identify underlying components. Components were subjected to sentiment analysis and summative content analysis with emergent categorial development to articulate themes.

Results: Seven themes emerged related to toxic relationships: (1) contextualizing storytelling with lymph nodes, (2) toxic behavior and venting emotions, (3) abandonment and abuse following diagnosis, (4) toxic relationships and social-related fears, (5) inner strength and navigating breast cancer over time, (6) assessing social relationships and interactions, and (7) community advice and support. Toxic relationships were commonly characterized by isolation, abandonment, and emotional abuse, which had profound emotional consequences for patients. Reddit facilitated anonymous venting about toxic relationships that helped patients cope with intense feelings and stress. Exchanging advice and support about navigating toxic relationships during breast cancer were core functions of the r/breastcancer community.

Conclusions: Findings emphasized the value of Reddit as a source of social support for patients with breast cancer experiencing toxic relationships. Clinicians who understand that many patients with breast cancer experience toxic relationships and considerable psychological sequelae are better prepared to support their patients’ holistic well-being. Further investigation of Reddit as a possible resource for advice, information, and support has the potential to help inform clinical practice and subsequently, patient health outcomes.

(JMIR Cancer 2024;10:e48860) doi:10.2196/48860

KEYWORDS
breast cancer; intimate partner violence; meaning extraction method; Reddit; sentiment analysis; social media; social support; toxic relationships; topic modelling
Introduction

Breast cancer has a way of making existing cracks in relationships even wider. Just like water will fill a crack in the road, freeze, and create a larger gap, breast cancer tends to permeate all parts of our lives and distance us from people with whom we have troubled relationships. [Original poster #85]

Projected rates of breast cancer in Canada have remained consistent over the past 5 years, with estimates that approximately 1 in 8 women will develop breast cancer in their lifetime and breast cancer will account for 25% of all new cancer cases [1-3]. Women’s experiences of breast cancer are influenced by the social determinants of health, particularly their social environment [4,5]. Among patients with breast cancer, strong social relationships have been found to act as a buffer to stress [6] and help to improve treatment effectiveness, psychological functioning, coping, survival, and quality of life, as well as prevent cancer recurrence [7-9]. Conversely, weak or nonexistent social relationships have been broadly linked to long-term psychological distress [10] and an increased risk of breast cancer progression, recurrence, and mortality [11,12]. However, there is a need for research that explores connections between social relationships and breast cancer outcomes among diverse populations and social contexts.

Intimate partners (eg, spouses and significant others) and immediate family members (eg, parents and siblings) are perceived as the most important social supports for patients with breast cancer [13,14], as they provide essential social-emotional, tangible, affection, and positive social interaction support [15]. For example, partners commonly serve as the primary caregivers of patients with cancer [16]. However, not all social relationships are supportive [17]. Patients who experience intimate partner violence (IPV) may face a lack of support due to the abusive behaviors of their partner [18]. IPV, understood as physical, psychological, or sexual abuse within the context of coercive control by an intimate partner [19], concurrently affects an estimated 12.5% of patients with breast cancer [20]—and this is likely to be an underestimation given underreporting of IPV [21]. Similarly, patients may be negatively affected by an unsupportive (but not necessarily abusive) partner [22], as well as abusive or unsupportive family members [23,24]. Aside from the patients themselves, immediate female family members are often most affected by a breast cancer diagnosis; unsupportive reactions often include being in denial about the diagnosis and abandoning the patient [25].

Toxic relationships are characterized by conflict, competition, undermining, disrespect, and a lack of cohesiveness [26]. Toxic relationships encompass unsupportive and abusive dynamics in both romantic (eg, a partner) and platonic (eg, a family member) contexts and are associated with emotional distress [26], which imparts numerous downstream mental and physical health consequences [27]. To compensate for unmet support needs, patients with breast cancer may expand their social networks via the internet, including social media [28]. Online forums are a popular means of accessing information and support related to breast cancer awareness, literacy, and treatment [29-33]. The use of online breast cancer forums grew exponentially between 2006 and 2010, growing from an estimated 282,000 new posts per year to over 1,270,000 new posts per year [34] and continues to increase over a decade later [35,36]. Despite data availability and the potential for knowledge advancement [33], research on patient social media use, particularly in the context of toxic relationships, is underexplored.

Reddit, the world’s third most popular social media platform, is an online forum dedicated to community-building, news dissemination, and discussion facilitation [37]. The Reddit platform consists of topic-specific subreddits (ie, forums), where all content is user-generated. Users subscribe to subreddits that interest them to see more related content. Users can post content, as well as comment and vote on others’ content. To join Reddit, users create a username and password—no identifiable information is required. Reddit’s capacity for anonymous participation and long-form, conversational content makes the platform a rich source of self-reported textual data [38]. The Reddit platform includes breast cancer–specific spaces that offer access to psychosocial support (eg, r/breastcancer), presenting a unique and valuable opportunity to explore how patients with breast cancer navigate toxic relationships after diagnosis. Previous research has provided preliminary insights into how patients with breast cancer use Reddit [39], but there is a notable gap in the literature regarding how patients with breast cancer describe toxic relationships and their psychosocial impacts on Reddit. Studying social media data has the potential to generate significant advances in knowledge [33], which can inform improvements to psychosocial support for patients with breast cancer experiencing toxic relationships and enhance care providers’ ability to promote patient well-being. Accordingly, this study sought to explore how people with breast cancer describe toxic relationships with their partners and immediate family members on Reddit.

Methods

Design

This exploratory, cross-sectional, topic-modeling study was conducted from December 2022 to February 2023 and aimed to explore how patients with breast cancer describe toxic relationships with their partners and immediate family members on Reddit. As of February 2023, the public r/breastcancer subreddit, established in 2011, included 13,900 subscribers and self-identified as a support and information group for people who have been diagnosed with breast cancer and their caregivers and loved ones. While Reddit generally attracts young White men of high socioeconomic status [38], demographics vary by subreddit and r/breastcancer is hypothesized to be largely composed of women [40].

Ethical Considerations

This study was deemed exempt from oversight by the author’s institutional ethics review board because all data were gathered from the public domain (per Article 2.2 of the Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans). The subreddit at the center of this study was public at the time of data collection and writing, meaning that any person could...
access its content at any time. It was therefore determined that r/breastcancer users had no expectation of privacy, negating the need for oversight by an ethics review board.

**Data Collection**

This subreddit was scraped for textual data from posts and comments using the Python Reddit application programming interface wrapper. No date limits were imposed. An iterative approach to keyword-based searching extracted posts (n=187) related to toxic relationships with partners and immediate family members. Two keyword strings were combined to scrape data: String 1 included words associated with a toxic relationship (eg, narcissist, boundaries, abuse, violence, assault, neglect, cheater, affair, divorce, toxic, abandon, and manipulate) and string 2 included words that identified people of interest in the immediate family of the user (eg, abuser, spouse, partner, marriage, significant other, parent, and sibling). To be scraped, posts were required to include a minimum of 1 keyword from both string 1 and string 2 (see Textbox 1).

Textbox 1. Keywords included in the final iteration of the search strategy.

<table>
<thead>
<tr>
<th>String 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>narcissist; boundaries; abuse; abusive; abusing; abused; violent; assault; assaulting; harassed; abusive; lie; lied; neglect; unsupportive; not supportive; not supporting; no support; cheater; cheated; cheating; affair; divorce; break up; breaking up; broke up; toxic; abandoned; manipulate; manipulated; emotionally unavailable; disown; alone; selfish; strained</td>
</tr>
<tr>
<td>String 2</td>
</tr>
<tr>
<td>abuser; husband; wife; partner; hubby; marriage; girlfriend; boyfriend; gf; bf; SO; significant other; spouse; mom; mother; mum; dad; father; parent; parents; sibling; sis; sister; brother</td>
</tr>
</tbody>
</table>

The scraped posts were then screened for eligibility by one of the authors (CAD), such that posts were ineligible if they addressed anyone other than a partner or immediate family member, were of an administrative nature posted by a moderator, were posted by a user who did not have breast cancer, aimed to exclusively seek or share medical information, or described toxic relationships outside of the context of breast cancer. After screening, 36 posts were eligible for inclusion. Eligible posts were scraped for comments (n=601), of which 98 were relevant (as determined by CAD using the eligibility criteria described above used to filter posts). Textual data were compiled into packets, where 1 packet represented the total relevant contributions (ie, posts and comments) from a single user, with an average of 260 words per packet. The final data corpus included 96 unique users with 36 posts and 98 comments (see Figure 1).

Figure 1. Final data corpus diagram.
mom has never bothered checking on me” could become “My mother doesn’t ever ask how I’m doing.”

**Analysis**

**Multistaged Approach**

A 2021 systematic analysis by Proferes et al [39] identified that computational-driven textual analysis (which includes topic modeling) was the primary means of knowledge generation using Reddit data. However, the authors also identified that such analyses are enhanced by the addition of qualitative and mixed methods analyses that account for contextual details [39]. Accordingly, the data corpus was subjected to a 3-stage, mixed methods analysis that used (1) the meaning extraction method (MEM), (2) qualitative sentiment analysis, and (3) summative content analysis (Figure 2).

**Stage 1: Meaning Extraction Method**

The MEM is a form of topic modeling useful for social media data exploration [43] and for generating large sample sizes of participants that are traditionally difficult to recruit [38]. Within other breast cancer–related research using Reddit data, the MEM has been described as a cost-effective means of identifying common themes described by patients [39]. The results of this method have been found to be similar in content and utility to those of traditional research methods in this domain (focus groups) [39].

The MEM identifies word clusters that co-occur in a data corpus, providing an efficient means of extracting meaningful patterns in language within high volumes of natural language data [38,43]. The Meaning Extraction Helper developed by Boyd [44] was used to analyze the textual packet data corpus, inclusive of the removal of common closed and open class words (<7.5%) and content word retention (≥5%), producing a binary output of each retained content word per OP (eg, 0=absent and 1=present). Boyd [45] also developed an open-access script for the R open-access statistical software (R Foundation for Statistical Computing), which was adapted for a principal component analysis (using a varimax rotation [43]). This produced a 9-component model that was considered acceptable ($\chi^2=3357.40, df=304, P<2.2e–16$, and KMO=0.538 [43]). Using a scree plot analysis, components 1 to 7 were retained. The 7 retained components explained 84.16% of the variance—a high proportion for a natural language application [43]. A high loading threshold of 0.50 was imposed on the content words within each component to promote thematic clarity and reduce cross-loading. The final components with refined content words were considered sufficiently strong (≥3 content words per component [46]).

**Stage 2: Qualitative Sentiment Analysis**

Qualitative sentiment analysis aims to assess the affective valence of components and their content words [47]. Modern qualitative sentiment analysis (ie, internet-based) is an increasingly popular and effective method of interpreting user-generated social media content [48]. Using the suzhet R package [49] and Afinn sentiment lexicon of –5 (negative sentiment) to 5 (positive sentiment [50]), a total model and 7 component-specific sentiment scores were computed based on content words. The content word tchp was changed to combination drug cancer therapy for the algorithm because it cannot assess acronyms.

**Stage 3: Summative Content Analysis**

Summative content analysis with emergent categorical development was used to articulate patterns and themes within the textual data packets for each of the 7 refined components [51]. Component categories (referred to as themes) were...
inductively developed to describe their overall message, inclusive of the use of sentiment scores to contextualize positioning. The 6-step approach to trustworthy thematic analysis by Nowell et al [52], rooted in the trustworthiness theory of Lincoln and Guba [53,54] was adopted.

Summary of Analysis

Quantitative topic modeling (with MEM) was combined with qualitative sentiment and content analysis to produce a comprehensive analytical framework capable of providing an overall interpretive assessment of the data corpus. The r/breastcancer subreddit includes thousands of textual data sources, requiring the combination of complex methods to efficiently target and isolate meaningful, manageable patterns from the large volume of natural language data [38,43]. The MEM is a computational method specifically developed to facilitate efficient filtering of large textual data sets, however, a second stage of qualitative or mixed methods–based analysis is recommended to facilitate deeper exploration and interpretation in context [39]. Accordingly, sentiment analysis was applied within MEM-generated principal components to facilitate the assessment and incorporation of considerations of user’s emotions and situational contexts. Following MEM and sentiment analysis, content analysis was used to deeply explore principal components through the lens of their socioemotional contexts to enrich interpretation and understanding. In sum, this combined mixed methods framework aimed to produce holistic, contextualized insights from MEM-generated categories, which is well-suited to complex, dynamic social media data.

Results

Overview of Themes

Seven distinct but related themes emerged from descriptions of toxic relationships by patients with breast cancer on Reddit, presented in order of explained variance proportion (highest to lowest) as follows: (1) contextualizing storytelling with lymph nodes, (2) toxic behavior and venting emotions, (3) abandonment and abuse following diagnosis, (4) toxic relationships and social-related fears, (5) inner strength and navigating breast cancer over time, (6) assessing social relationships and interactions, and (7) community advice and support. The overall corpus sentiment score was –4, indicative of very negative sentiment. Theme-specific sentiment scores reflect the average valence of retained content words within each component.

Theme 1: Contextualizing Storytelling With Lymph Nodes

*I’ll have to get my lymph nodes removed next, among other things. Treatment is lonely and miserable.* [OP 2]

The first theme was classified as neutral (\(\bar{r} =0.00\)) and included lymph, node, and pick as key content words. Lymph nodes functioned as context indicators in users’ stories about toxic relationships to highlight their temporality within cancer treatment. For example, one user was undergoing chemotherapy while navigating a toxic relationship with their mother. This OP prefaced their post by sharing, “After a lot of treatment, my cancer went from grade 3 to grade 1. My lymph nodes shrunk as well” (OP 66).

They then went on to disclose unsupportive behavior from their mother, stating, “My mom doesn’t think I’m capable of making my own decisions—but I am. I’ve picked excellent physicians and made it to all of my appointments” (OP 66).

Theme 2: Toxic Behavior and Venting Emotions

*I’m going to vent because I think it’s better to write than to cry...* [OP 65]

The second theme was classified as neutral (\(\bar{r} =0.00\)) and described toxic relationships that the user experienced a strong emotional reaction to, which prompted them to vent their emotions on Reddit. Key content words included boundary, effort, vent, upset, and stress. Users reported a variety of toxic behaviors, such as boundary violations and disrespectful or abusive actions. Venting was commonly used to cope with powerful negative emotions associated with toxic relationships.

Users felt unsupported when their partners or families reacted to their diagnosis by becoming detached or distressed to the extent of relying on the patient for support. To illustrate, one user expressed disappointment in their father’s silence after diagnosis, stating, “My dad isn’t there for me. I guess I shouldn’t be surprised, he’s always been like this” (OP 45).

Other users were frustrated with bearing the emotional burden for others regarding their cancer. For example, one OP resented their husband for expecting them to manage his emotions, sharing, “I did my best to explain that I needed him to be my rock. He got upset... he wanted us to be mutually supportive. But he doesn’t have cancer... I do!” (OP 62).

Some OPs described being disrespected and emotionally abused following their diagnosis. For instance, one OP shared that their partner told them, “Lately, you aren’t sexually desirable to me without your natural breasts. I miss them and how they felt... probably even more than you do” (OP 65).

Similarly, another OP disclosed experiencing emotional and verbal abuse from their partner both before and after their breast cancer diagnosis. This OP shared feeling extremely upset that just 2 weeks after their diagnosis, their partner asked them, “How long are you going to pull the breast cancer card?” (OP 85).

Toxic relationships described within this theme were strongly associated with venting, that is, posting negative, emotionally charged content. For example, an OP trying to cope with being isolated by their family prefaced their story by writing, “Heads up that this is a massive, sad vent post. Sorry but I feel like I need to shout into the void” (OP 34).

Theme 3: Abandonment and Abuse Following Diagnosis

*Anyone else dealing with an emotionally abusive spouse before and during cancer? I’m trying to get away and he’s being awful.* [OP 40]
Theme 3 was classified as slightly negative (\( r = -1.00 \)) and captured how patients in toxic relationships were abandoned or emotionally abused by their partners following their diagnosis. Key content words included devastate, experience, and abuse. Patients who navigated abandonment or abuse concurrently with a breast cancer diagnosis reported feeling emotionally devastated.

Abandonment was especially common after disclosing a breast cancer diagnosis. For example, one OP shared that their husband abandoned them on the way home from their diagnosis appointment, stating, “He said he won’t look after the kids and plans on leaving” (OP 9). Other users were abandoned as treatment began. Many users who shared stories of abandonment described emotional whiplash, characterized by a sudden, unexpected transition from feeling secure in their relationship to feeling betrayed following abandonment. As illustrated by one user, “He made me feel cared for, loved, and safe... until I said I was considering a mastectomy. Then he shut me out” (OP 12).

The emotional impacts of betrayal were devastating. An OP whose long-term partner unexpectedly broke their promise to stick by them during treatment shared, “I am completely devastated. I am infuriated. He and my body betrayed me. I am so furious” (OP 86).

Of partners who stayed following a diagnosis, many subjected the patient to emotional abuse. One OP was told that they deserved their cancer, recounting, “He used my cancer against me by saying I got it because I’m weak and that’s just natural selection at work. He told me not to bother with treatment and to just let nature run its course” (OP 76).

Other experiences involved infidelity, threats of child abandonment, accusations of faking symptoms, and coercion in treatment choices. Emotional abuse was repeatedly described as devastating. For example, an OP whose spouse had been emotionally abusive for years posted, “What can I do to stop feeling devastated that my husband feels I should be punished all the time?” (OP 40).

**Theme 4: Toxic Relationships and Social-Related Fears**

*Do any of you also feel like the emotional consequences of breast cancer are almost worse to deal with than the physical? [OP 66]*

The fourth theme was classified as slightly negative (\( r = -0.75 \)) and focused on social-related fears associated with breast cancer. Key content words included biopsy, tchp, and scare. Patients’ fear stemmed from anticipating or experiencing a negative reaction to their breast cancer by a toxic family member or partner. For example, an OP who disclosed a toxic family shared dredging their reaction to their cancer, expressing, “The fear of how my family will react to my breast cancer diagnosis is nearly as overwhelming as the actual diagnosis” (OP 81).

Other users felt scared because they had already experienced an unsupportive reaction by a toxic family member or partner to their cancer. For example, one OP felt scared and hopeless after being gaslit by their partner about their diagnosis, sharing, “He was trying to tell me that my breast cancer was all in my head, despite having seen my biopsy results and meeting with multiple members of my medical team” (OP 3). Similarly, an OP whose family neglected to support them after learning of their diagnosis expressed, “My family doesn’t care about me or my breast cancer. It makes me feel scared and alone” (OP 34).

**Theme 5: Inner Strength and Navigating Breast Cancer Over Time**

*I thought to myself that if my cancer ever came back, I’d rather deal with it alone than with a person like that. [OP 32]*

Theme 5 was classified as slightly positive (\( r = 1.00 \)) and highlighted how breast cancer was disruptive to the lives of patients. Key content words included future, matter, and strength. Users described how health and social adversity influenced their inner strength. Toxic relationships that emerged after diagnosis were especially trying for patients. For instance, one OP expected their partner’s support as they began cancer treatment (as their partner had promised). However, the OP’s partner abruptly took back their commitment, leaving the OP to navigate cancer alone: “They sent me a message the next day and said they don’t want anything to do with me” (OP 12).

Inner strength emerged as a dynamic construct that was both challenged by experiencing a breast cancer diagnosis and toxic relationships and enhanced by surviving these adverse experiences. Many users believed that surviving breast cancer concurrently with exposure to toxic relationships was a testament to their inner strength. For example, one OP attributed their inner strength to recovering from breast cancer while navigating a lack of empathy and support from their spouse. This OP stated, “I feel 100% confident that I am a strong, intelligent woman who can face almost anything” (OP 64), while sharing that they had received a new cancer diagnosis. Inner strength also enabled users to regain a sense of control over how they were going to navigate living with a breast cancer diagnosis. For example, an OP who was abandoned by their partner after being diagnosed stated, “I finally felt strong enough to delete his contact information because I couldn’t stop myself from calling him–it was the best choice I could have made” (OP 47).

**Theme 6: Assessing Social Relationships and Interactions**

*I am immensely grateful for you all for helping me navigate a chaotic and frustrating moment. [OP 85]*

The sixth theme was classified as marginally positive (\( r = 0.20 \)) and described how OPs assessed their social relationships and interactions. Key content words included conversation, response, listen, regret, and grateful. Users assessed the quality of social support from family based on whether they felt judged, subjected to toxic positivity, or made to listen to unsolicited advice. For example, an OP with an emotionally unsupportive family shared, “I think a lot of family think it’s helpful when they shove positivity down our throats. What we really need is support and someone to listen without trying to solve all our problems” (OP 82).
For some users, responses to breast cancer unveiled toxic relationships that they regretted having to face. For example, an OP with unsupportive parents shared, “I regret that my breast cancer forced me to confront that my parents never have and still don’t support me how I need them to” (OP 4). However, OPs who discovered both toxic and supportive relationships during cancer expressed gratitude for the sources of support they did have. As one OP stated, “Sometimes I get jealous of people whose parents love and support them, but then I remember the rest of my friends and family who showed up for me when I needed them, and I’m grateful” (OP 32). The subreddit community was repeatedly praised by users because it was such a valuable source of support. For instance, one OP shared, “I am endlessly grateful for the knowledge and resilience of this community” (OP 53).

Theme 7: Community Advice and Support

I know what it feels like to be abandoned. I could tell you all the red flags in a man's behavior... but just trust me—it’s better to be alone. You dodged a MASSIVE bullet. A person who lacks compassion about your breast cancer is NOT a good life partner. Please message me if you need someone to vent to. I really do understand... and you’ve got this. [OP 89]

The seventh theme was classified as marginally negative (r = -0.17) and characterized a core function of r/breastcancer: providing advice and support. Key content words included money, quit, and follow. The subreddit facilitated advice regarding various topics, especially related to navigating financial matters and treatment options in the context of a toxic relationship.

Numerous users offered money-related advice to OPs facing difficult financial situations because of toxic relationships. Situations included financial coercion, exploitation, and manipulation following cancer disclosure and managing finances during separation from a toxic partner. For instance, one OP was abandoned by their partner during a joint real estate purchase. A community member with self-professed real estate expertise strongly advised the OP against continuing with the investment, writing, “I’m begging you... please DO NOT sign anything else! Lose your money... that’s not important... please do not continue with this purchase” (OP 89).

Members also counseled OPs about postmastectomy reconstruction by offering advice on how to reduce social pressure and prioritize personal preferences. For example, one OP shared how they resisted their partner’s pressure to follow reconstruction, stating, “I made him look at photos of reconstruction to show him that it’s not a free boob job and can be ugly. He changed his tune real quick” (OP 39).

Members who were ultimately pressured into reconstruction strongly encouraged OPs to follow their instincts. For example, one member who was coerced into reconstruction by their husband advised, “I constantly wish I went flat instead. If I had to do it again I would listen to my gut and go flat” (OP 8).

Similarly, it was common to share advice about treatment adherence. Many OPs struggling with a lack of support expressed wanting to quit treatment. While members empathized with users and understood their feelings, they ultimately encouraged continuing. For example, one OP shared, “I’m just sick of this. I’m pretty sure I’m done with it all” (OP 34).

The community offered empathy, such as, “When I was in the middle of your treatment, I was frustrated too and tried to quit every week” (OP 7), as well as advice, for example, “Don’t stop treatment without a good reason. It’s a gift in spite of tough side effects because it keeps us alive” (OP 48).

Discussion

Principal Results

This study explored the use of the r/breastcancer subreddit by patients to describe toxic relationships with their partners and immediate family members. Themes highlighted patients’ lived experiences of toxic relationships, emotional impacts, and support from the subreddit community. A key finding was that many people with breast cancer sought out the r/breastcancer subreddit to share their experiences of toxic relationships, often including descriptions of abandonment, isolation, and emotional abuse within this context. Further, this study presented compelling evidence that toxic relationships impart profound emotional consequences for patients and that some patients cope with these strong emotions through online venting. This work also emphasized the value of online communities like Reddit as alternative, complementary sources of support for patients experiencing toxic relationships.

Comparison With Prior Work

Abandonment and Betrayal as Common Experiences

These findings suggest that abandonment is a common experience for patients with breast cancer following diagnosis. Prior research has lacked consensus regarding the risk of abandonment among patients with breast cancer after diagnosis [55,56]. Generally, however, women are more likely to be abandoned by a partner after being diagnosed with a serious medical illness [57]. Further, distancing is the most prevalent unsupportive response experienced by a patient following their breast cancer diagnosis [23]. Fears and feelings of abandonment following diagnosis are also well-documented within breast cancer research [58-60]. Given this understanding, and considering that Reddit data can be regarded as an authentic representation of user experiences [61], it is reasonable to conclude that these findings are suggestive of an increased risk of abandonment for patients with breast cancer.

A novel finding was the occurrence of emotional whiplash, where a patient was initially promised support by their partner but was later abandoned unexpectedly. The emotional transition from security to betrayal was repeatedly reported as devastating. There is limited research describing betrayal in the context of abandonment and breast cancer, but it is known that feelings of betrayal in this context can reduce the desire for future relationships [62]. Broadly, the loss, disruption, and deterioration of social ties are some of the most stressful experiences a patient with cancer can face [6], which makes abandonment a serious risk factor for reduced mental health [63]. Comprehensive cancer care entails stress-reducing psychosocial interventions [63], but
a limited understanding of the psychological effects of betrayal hinders clinicians’ ability to optimally manage abandonment-related stress.

**Anonymous Venting Enables Disclosure of Toxic Relationships**

The central role of venting within the r/breastcancer community highlighted the unique socioemotional needs of patients with breast cancer in the context of toxic relationships. Toxic relationships impart emotional consequences that can be difficult to navigate and cope with [26]. Venting is a disinhibitory, emotion-focused strategy for coping with stress [64,65]. Venting can be considered a form of expressive writing, that is, writing that describes a deeply personal experience [66], which is well-evidenced to facilitate coping with psychological distress [67]. Online venting was consistently described as cathartic among patients in this study, aligning with prior evidence of patients with breast cancer seeking support in online communities during periods of stress [68] and perceiving reduced stress after they vent online [69]. Further, patients with breast cancer who self-manage their emotions by narrating their experiences are known to experience strong psychological benefits [70].

It might be expected that the stigma attached to breast cancer and toxic relationships would hinder disclosure [58,71], however, seeking out group-oriented support is reportedly most common for diseases considered stigmatizing [72]. The latter position is consistent with this study, as venting posts often included stigmatized thoughts and feelings (eg, wanting to ‘give in’ to cancer or discussing abuse without wanting to leave the relationship). Further, it appeared that Reddit’s capacity for anonymity created a sense of safety that made patients comfortable disclosing information considered stigmatizing, which is consistent with existing evidence [73]. Overall, patients appeared to perceive anonymous venting via Reddit as an effective, safe strategy for coping with stress from toxic relationships. Interventions that aim to promote coping among this patient population would likely benefit from integrating anonymity to encourage uninhibited self-expression.

**Advice About Navigating Toxic Relationships**

Validating the feelings of other users, as well as soliciting and providing advice regarding toxic relationships, were core activities within r/breastcancer. It was previously known that participation in online forums contributes to the practical, informative, and emotional empowerment of patients with breast cancer [74]. However, this study uniquely identified that community members on Reddit often urged OPs to leave or go against the wishes of their abusive partner. While well-intentioned, this advice may not always be safe or practical. Leaving an abusive partner can be the most dangerous time in the relationship due to an increased risk of retaliation [75]. Similarly, acting in a manner that might antagonize an abuser can initiate or escalate relationship discord and consequently increase the risk of violence [76]. Furthermore, patients who depend on an abusive partner (eg, for caregiving, access to health insurance, and transportation to appointments [77]) may be unable to leave or risk the relationship by acting defiantly [78]. Resultantly, relationship advice received on Reddit by patients with abusive partners may have been incompatible with their reality or suboptimal in promoting their safety.

This indicates a knowledge gap concerning safety planning within r/breastcancer; safety planning can be understood as the development of strategies to reduce the risk of abuse and enhance support [79]. Safety planning is a proven, widely endorsed health promotion intervention that is effective both within an abusive relationship and after leaving [80,81]. Considering the prevalence of abuse among patients with breast cancer [20] and that many seek support in online forums such as Reddit [39], it could be useful to raise awareness of safety planning within r/breastcancer as a health promotion strategy. Further, considering the importance attributed to inner strength by patients in this study, building awareness of strengths-based approaches to safety planning [82] could be particularly useful. For example, community moderators could pin relationship-related resources (eg, hotlines and informative websites) as the top comment under posts about challenging, potentially toxic relationships. However, a needs assessment would be best suited to developing an IPV-related intervention considered acceptable and effective within r/breastcancer.

**Clinical Implications**

Psycho-oncology care teams play a critical role in optimizing health outcomes for patients with breast cancer, yet the emotional well-being of patients with cancer is often underreported and underexplored [83]. Patient-reported social media data offers real-time insights into patient experiences and needs which can be beneficial for informing clinical practice [33,83].

Clinicians who understand that many of their patients with breast cancer are negatively affected by toxic relationships are better prepared to support their emotional well-being. Acquiring knowledge about practices and resources that foster coping and inner strength, including venting and safety planning, can contribute to improved patient outcomes.

Some clinicians may be unfamiliar with the advantages of online forums for patients, but recognizing the potential benefits could enhance care [84]. Recommending Reddit as a possible source of advice, information, and support could be a valuable addition to clinical practice for patients navigating breast cancer and toxic relationships. However, because digital literacy is often overlooked in breast cancer care [85], clinicians who concurrently promote digital literacy can empower their patients to access online communities and ultimately, improve their health outcomes.

**Limitations**

There are limitations to this work. First, the analysis was conducted by a single researcher, which may have introduced bias in data interpretation. The analysis also relied heavily on automated methods that may have been inadequate in fully capturing nuance or interpreting context cues in textual data. Second, these data are self-reported, which may have resulted in biased perspectives. While users in this sample self-identified as patients with breast cancer, it was not possible to validate this. These data may have inadvertently included content from online robots or people without breast cancer, and thus may not
accurately reflect the experiences of the target population. Additionally, these data were scraped from a single social media platform and may not be representative of the experiences of patients who use other social media platforms, do not use Reddit to discuss their personal lives, or lack access to an internet-enabled device. No demographic information was available to further contextualize findings. It is important to note that these results only relate to experiences of emotional abuse, as physical and sexual abuse were not represented in the data. Furthermore, all participants could write in English, were digitally literate, and had access to the internet, meaning that the findings may not represent the experiences of patients who are nonanglophone or lack technological access or literacy. Caution should be used when applying these findings to other patients with breast cancer.

Conclusions
This study identified that toxic relationships described by patients with breast cancer on Reddit were common and characterized by abandonment, abuse, and unsupportive behaviors. Patients often experienced profound emotional reactions to this form of social stress and anonymous venting on Reddit was described as an effective coping mechanism. Some patients described breast cancer and toxic relationships as adverse experiences that ultimately enhanced their inner strength. Overall, the r/breastcancer community appeared to be a means of exchanging advice, information, and support for patients experiencing toxic relationships. Clinicians who understand that their patients may be negatively affected by toxic relationships are better prepared to support their holistic well-being. Further investigation of Reddit as a possible source of advice, information, and support has the potential to help inform clinical practice and subsequently, improve patient health outcomes.

Data Availability
The data sets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors’ Contributions
CAD was involved in the conceptualization, data curation, formal analysis, investigation, methodology, and both original draft preparation and review and editing of the manuscript. RB, KTJ, and TM contributed to the conceptualization, supervision, methodology, and review and editing of the manuscript.

Conflicts of Interest
None declared.

References


42. Boyd RL. MEH: Meaning Extraction Helper (Version 2.3.02). Ryanboid.io. URL: https://www.ryanboid.io/software/meh [accessed 2023-02-10]


60. Yañez-Jiménez MD, Lucero JD, del Carmen de la Trinidad CAMG. Information needs about cancer treatment, fertility, and pregnancy: qualitative descriptive study of reddit threads. JMIR Cancer 2020;6(2):e17771 [Full text] [doi: 10.2196/17771] [Medline: 33263547]


Abbreviations

IPV: intimate partner violence
MEM: meaning extraction method
OP: original poster
Exploring Web-Based Information and Resources That Support Adolescents and Young Adults With Cancer to Resume Study and Work: Environmental Scan Study

Clarissa E Schilstra¹, BA, PhD; Sarah J Ellis¹, DClinPsy, PhD; Jennifer Cohen²,³, MSc, PhD; Alana Gall⁴, MSc, PhD; Abbey Diaz⁵, MSc, PhD; Kristina Clarke³, BCom, PGDipSci; Gadiel Dumlao¹, BSc; Jennifer Chard⁶, MBBS; Therese M Cumming⁷,⁸, MEd, PhD; Esther Davis¹, BSc, PhD; Haryana Dhillon⁹, BSc, MA, PhD; Mary Anne Burns⁹, BSc; Kimberley Docking¹⁰, BPath, MSPA, CPSP, PhD; Eng-Siew Koh¹¹,¹²,¹³, MBBS; Josephine O'Reilly¹⁴; Ursula M Sansom-Daly¹,¹⁵, BSc, PhD, MAPS; Joanne Shaw⁹, BAplSc, BPsych, PhD; Nicole Speers¹⁴; Natalie Taylor², PhD; Anthea Warne²; Joanna E Fardell¹,⁶, BSc, M Clin Neuropsych, PhD

¹Behavioural Sciences Unit, Kids Cancer Centre, Sydney Children's Hospital, Faculty of Medicine and Health, Randwick Clinical Campus, University of New South Wales Sydney, Randwick, Australia
²Faculty of Medicine and Health, Randwick Clinical Campus, University of New South Wales Sydney, Kensington, Australia
³Canteen Australia, Newtown, Australia
⁴National Centre for Naturopathic Medicine, Faculty of Health, Southern Cross University, Lismore, Australia
⁵School of Public Health, Faculty of Medicine, University of Queensland, Brisbane, Australia
⁶Western Sydney Youth Cancer Service, Crown Princess Mary Cancer Centre, Westmead Hospital, Westmead, Australia
⁷Faculty of Arts, Design and Architecture, University of New South Wales Sydney, Kensington, Australia
⁸Disability Innovation Institute, University of New South Wales Sydney, Kensington, Australia
⁹School of Psychology, Faculty of Science, University of Sydney, Camperdown, Australia
¹⁰Faculty of Medicine and Health, University of Sydney, Camperdown, Australia
¹¹South West Sydney Clinical School, Faculty of Medicine and Health, University of New South Wales Sydney, Liverpool, Australia
¹²Liverpool and Macarthur Cancer Therapy Centres, Liverpool, Australia
¹³Ingham Institute for Applied Medical Research, Liverpool, Australia
¹⁴Cancer survivor representative, New South Wales, Australia
¹⁵Sydney Youth Cancer Service, Nelune Comprehensive Cancer Centre, Prince of Wales Hospital, Randwick, Australia

Corresponding Author:
Clarissa E Schilstra, BA, PhD
Behavioural Sciences Unit, Kids Cancer Centre, Sydney Children's Hospital
Faculty of Medicine and Health, Randwick Clinical Campus
University of New South Wales Sydney
Behavioural Sciences Unit, Kids Cancer Centre
Level 1 South, Sydney Children's Hospital, High St
Randwick, 2031
Australia
Phone: 61 405679809
Email: c.schilstra@unsw.edu.au

Abstract

Background: Adolescents and young adults (AYAs) diagnosed with cancer experience physical, cognitive, and psychosocial effects from cancer treatment that can negatively affect their ability to remain engaged in education or work through cancer treatment and in the long term. Disengagement from education or work can have lasting implications for AYAs’ financial independence, psychosocial well-being, and quality of life. Australian AYAs with cancer lack access to adequate specialist support for their education and work needs and report a preference for web-based support that they can access from anywhere, in their own time. However, it remains unclear what web-based resources exist that are tailored to support AYAs with cancer in reaching their educational or work goals.
Objective: This study aimed to determine what web-based resources exist for Australian AYAs with cancer to (1) support return to education or work and (2) identify the degree to which existing resources are age-specific, cancer-specific, culturally inclusive, and evidence-based; are co-designed with AYAs; use age-appropriate language; and are easy to find.

Methods: We conducted an environmental scan by searching Google with English search terms in August 2022 to identify information resources about employment and education for AYAs ever diagnosed with cancer. Data extraction was conducted in Microsoft Excel, and the following were assessed: understandability and actionability (using the Patient Education and Materials Tool), readability (using the Sydney Health Literacy Laboratory Health Literacy Editor), and whether the resource was easy to locate, evidence-based, co-designed with AYAs, and culturally inclusive of Aboriginal and Torres Strait Islander peoples. The latter was assessed using 7 criteria previously developed by members of the research team.

Results: We identified 24 web-based resources, comprising 22 written text resources and 12 video resources. Most resources (21/24, 88%) were published by nongovernmental organizations in Australia, Canada, the United States, and the United Kingdom. A total of 7 resources focused on education, 8 focused on work, and 9 focused on both education and work. The evaluation of resources demonstrated poor understandability and actionability. Resources were rarely evidence-based or co-designed by AYAs, difficult to locate on the internet, and largely not inclusive of Aboriginal and Torres Strait Islander populations.

Conclusions: Although web-based resources for AYAs with cancer are often available through the websites of hospitals or nongovernmental organizations, this environmental scan suggests they would benefit from more evidence-based and actionable resources that are available in multiple formats (eg, text and audio-visual) and tailored to be age-appropriate and culturally inclusive.

(JMIR Cancer 2024;10:e47944) doi:10.2196/47944

KEYWORDS
adolescent; cancer; education; employment; information needs; oncology; online information; quality of life; resource; return to work; school; study; supportive resources; treatment; young adult

Introduction

A diagnosis of cancer in adolescence and the young adult years can lead to significant and long-lasting disruptions to key developmental milestones [1-3]. Adolescents and young adults (AYAs) with cancer are at risk of poor long-term medical and psychosocial outcomes due to delays in diagnosis and lagging improvements in survival rates compared with those diagnosed as children or adults [1]. Symptoms and late effects from cancer treatment can negatively impact AYAs’ education and work engagement [4-7]. A weakened immune system, nausea, fatigue, neuropathy, poor cognitive functioning, poor social well-being, and mental health challenges such as depression and social anxiety are just some of the physical symptoms and late effects AYAs must contend with [3-9]. Challenges with education and work engagement may also drive distressing symptoms such as poor social well-being and depression [3,9,10].

Poorer outcomes are exacerbated by the limited age-appropriate services targeting their unique needs [11]. During adolescence and the young adult years, broadly defined as the ages of 15-39 years, young people are expected to participate in and complete education and training, obtain employment, and achieve financial independence, all while navigating social and intimate relationships to develop and evolve their identity [12]. A cancer diagnosis and treatment during adolescence and the young adult years can interrupt or delay these developmental tasks [12].

Extended absences and difficulties engaging with education or work are common for survivors of AYA cancer and can have a lifelong negative impact on AYAs’ educational and work goals, quality of life, and psychosocial and financial well-being [2]. Survivors of AYA cancer miss significantly more days of school than their peers [11]. AYAs may miss 40-60 days of school within the first year following their diagnosis [13], and Australian data suggest nearly 50% of survivors of AYA cancer have not fully returned to education or work up to 24 months post diagnosis [14]. Nearly 40% of AYAs report their employment goals were negatively affected by cancer [15], and survivors of AYA cancer are more likely than peers without a history of cancer to report an increased number of missed workdays as a result of illness or disability (11.9% of survivors vs 6.7% of controls) [16]. Survivors of AYA cancer are also more likely than peers without a history of cancer to report employment disability (being unable to carry out employment or work requirements at all or needing to do so with disability provisions; 34.1% of survivors vs 23.9% of controls) [16]. However, qualitative studies suggest that survivors of AYA cancer experience trouble navigating public support, education, and employment systems, which puts them at a disadvantage by contributing to increased financial hardship, fear, and uncertainty around their education and employment situations. Financial hardship in itself serves as a barrier to AYAs achieving their education goals by making it difficult for AYAs to afford education, particularly at a university level, or to keep up with repayment of education-related debts [17].

In contrast, AYAs who are able to remain more engaged with their education or work report decreased psychological distress [18] and improved social well-being [1,19]. Yet, few interventions supporting AYAs to remain engaged with education or work have been systematically evaluated [2]. The only such service evaluated in Australia is an educational and vocational counseling service based in a major cancer center in Victoria, Australia, which provides in-depth, tailored support through trained advisors to AYAs diagnosed with cancer between the ages of 15 and 25 years [20].

To date, 209 AYAs have received support through this program.
73% of AYAs were able to engage in education or vocation or were receiving support through an external source [20]. Access to such personalized programs may be limited due to cancer centers’ resources and AYAs’ distance to their treatment centers [21,22]. Individualized consultation requires synchronous engagement, usually during school or work hours, which may further prevent uptake for AYAs who have some level of participation in education or work [23].

The average Australian AYA spends approximately 14 hours per week on the internet, and survivors of AYA cancer are highly engaged with social media as well as web-based cancer resources [24,25]. Many Australians with cancer living in rural or remote locations rely on web-based resources to navigate the impacts of a cancer diagnosis and its treatment on their education or career [26,27]. Given this reliance on web-based information, it is critical to consider the extent to which web-based information resources are equitably accessible by Australian AYAs. Previous studies and reviews have defined equity of access to web-based information as the provision of web-based information that is easy to find, provided in a range of formats (eg, text, video, and audio), understandable or readable for individuals with varying abilities and health literacy levels, and culturally and linguistically inclusive [28-30]. The importance of equitable access to health information for Australians is paramount, considering how many people live in rural locations and the cultural and linguistic diversity of the country. Approximately 3.2% of the population identify as Aboriginal and Torres Strait Islander peoples, 30% were born overseas, and 21% of families speak a language other than English at home [31].

Ensuring equitable access to web-based information resources requires consultation with target populations (ie, co-design of information resources with a culturally and linguistically diverse group of AYAs diagnosed with cancer), as well as a focus on providing information that is evidence-based. However, no previous research has assessed what web-based information resources exist to support Australian survivors of AYA cancer in their engagement with education or work, and to what extent resources are equitably accessible. Therefore, this study aimed to determine the following: (1) What web-based resources exist for engaging with education or work after a cancer diagnosis that AYAs with cancer are likely to encounter when conducting a Google search? and (2) Of the identified resources, to what degree are they understandable, actionable, readable, easy to locate, evidence-based, co-designed with survivors of AYA cancer, and culturally inclusive?

Methods

Overview

Web-based resources for AYAs with cancer are typically provided through hospitals and nongovernmental organizations (NGOs) rather than through academic journals or research databases. Therefore, we chose to conduct an environmental scan rather than a systematic review, using a standard search engine rather than academic databases. Environmental scans have demonstrated usability in identifying health information resources across a range of health disciplines [32,33]. While there is no consensus regarding optimal methods for conducting an environmental scan of health information resources, environmental scans take a higher-level approach than systematic reviews or qualitative evaluation studies to identify available resources, tabulate yes or no responses to whether web-based resources possess certain qualities, and determine the basic usability of resource content [33,34]. We opted to follow similar methods used by Ruble et al [32] in their 2019 publication assessing web-based resources to support children returning to school during or after cancer treatment and methods used by Schiffman et al [35] in their 2006 study on internet use among survivors of AYA cancer. In keeping with these previous studies, 2 researchers led the search and data extraction, and we used validated measures of understandability, actionability, and readability to conduct a basic assessment of available resources. We also tabulated whether resources were easy to locate, evidence-based, developed through co-design with AYAs, and culturally inclusive of Indigenous populations. We conducted structured searches through Google and extracted data in Microsoft Excel (Microsoft Corporation).

Consumer Involvement

Consumer involvement in the design of this study was central to our methods. Chief investigators included the researcher and clinician chief investigators in addition to 2 survivors of AYA cancer (authors CES and NS) and 1 parent of a survivor (author JO). Together, the chief investigator team met in November 2021, March 2022, and May 2022, to develop the environmental scan protocol, including search terms and methods.

Searches

We searched Google Australia with English search terms between August 8 and 19, 2022 (Textbox 1). No limits were applied to the country, as we wanted to replicate the way survivors of AYA cancer currently access information to support their return to education or work.

Textbox 1: Google search terms

- [Common search terms for AYA cancer]
- [Resources for education and career]
- [Survivor stories]
- [Support organizations]
- [Information for families]
- [Cancer survivorship]
- [AYA cancer]
- [Cancer education]
- [Career counseling]
- [Cancer support groups]
- [Ayurveda for cancer]
- [Cancer survivorship

https://cancer.jmir.org/2024/1/e47944
Search terms were created by combining search words from 2 themes. The first theme designated the target population of AYAs with or surviving cancer, and the second designated information and resource content related to “returning to study or work.” Using the words listed in Textbox 1, CES and GD independently conducted 24 unique searches combining the 2 groups of search terms with “AND” (eg, “adolescent cancer AND school”). Before and between each new search, the browser cache was reset. All searches were conducted from Sydney, New South Wales, Australia. However, to identify whether there may be any difference in search results based on location in Australia, GD also conducted 6 of the 24 searches using a virtual private network and changed the search location to Perth, Western Australia, which is located on the opposite side of Australia from Sydney.

Resource Selection
Typically, a Google search will present 10 results per page, meaning 50 results would be presented across 5 pages. Although the average internet user will only click on results appearing in the first 10 Google search results [36], we opted to maximize, the identification of relevant results by reviewing the first 50 results for eligibility (CES and GD) [37]. Eligible websites, documents, videos, and audio-visual resources were those that provided text-based information, video, or audio-visual information in English and were directed primarily toward AYAs returning to study (any level) or work after a cancer diagnosis. Websites, documents, and videos or audio-visual resources targeting parents or family members of AYAs were excluded. Academic papers, media stories, and blogs were also excluded.

Data Extraction
An Excel spreadsheet was developed to include drop-down menus to record key data (Table 1). Data extraction fields were partially based on a previous review of web-based resources conducted by Ruble et al [32] in 2020. CES and GD independently conducted data extraction and recorded the addresses of websites meeting eligibility criteria in separate Excel spreadsheets, reconciled their searches, and removed duplicates.
### Table 1. Data extraction fields and response options.

<table>
<thead>
<tr>
<th>Field</th>
<th>Response option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target audience</td>
<td>AYA-specific or other</td>
</tr>
<tr>
<td>Setting</td>
<td>Education (inclusive of secondary or tertiary) or employment</td>
</tr>
<tr>
<td>Source creator</td>
<td>• Nongovernmental organization</td>
</tr>
<tr>
<td></td>
<td>• Cooperative group or professional organization</td>
</tr>
<tr>
<td></td>
<td>• Health care institution</td>
</tr>
<tr>
<td></td>
<td>• State or federal government organizations</td>
</tr>
<tr>
<td></td>
<td>• Media publications</td>
</tr>
<tr>
<td>Country of origin</td>
<td>Country name</td>
</tr>
<tr>
<td>Access to website</td>
<td>Publicly available, subscription, or user profile</td>
</tr>
<tr>
<td>Purpose of website</td>
<td>Information, advertising, support, or intervention</td>
</tr>
<tr>
<td>Date of last review of information</td>
<td>Date</td>
</tr>
<tr>
<td>Cancer type</td>
<td>Diagnosis name, resource designed for AYAs with chronic illness more broadly but includes cancer, or no cancer type specified</td>
</tr>
<tr>
<td>Media used to convey information</td>
<td>Written descriptive text, video or YouTube, images, stories or vignettes, quotes from consumers, or other</td>
</tr>
<tr>
<td>Support, tools, or information provided by the resource</td>
<td>Checklists, letter templates, strategies, access to career or education counseling, support group or network, or other</td>
</tr>
<tr>
<td>Evidence-based</td>
<td>Yes or no</td>
</tr>
<tr>
<td>Co-designed</td>
<td>Yes, no, or unclear: describes consultation with survivors but not methods for this consultation</td>
</tr>
</tbody>
</table>

### Assessment of Resources and Data Synthesis

#### Was the Resource Understandable and Actionable?

For both text and audio-visual resources, we used the Patient Education Materials Assessment Tool (PEMAT) [38] to assess the understandability and actionability of the resources. Understandability refers to whether the meaning is comprehensible, taking multiple elements into account, such as word complexity and the layout or structure of the information [38]. Actionability refers to whether or not a resource provides content in a way that consumers can easily determine what they need to act on or do based on the content presented [38]. PEMAT for written materials consists of 17 items assessing understandability and 17 items assessing actionability, all of which are scored as agree, disagree, or unsure [38]. The PEMAT for audio-visual materials includes 13 items assessing understandability and 4 assessing actionability. The PEMAT generates percentage scores (0%-100%) which is the proportion of the responses assessed as having been met (agree). Scores of 100% indicate optimal understandability or actionability; scores of 70% indicate adequate understandability or actionability [38,39]. The PEMAT has been used previously in a review of information resources for students with cancer [32] and an evaluation of other web-based information for many illnesses, including cancer [37]. The PEMAT demonstrates good reliability and ease of use, with interrater reliability scores of 0.92 for understandability and 0.93 for actionability, and 92% of raters agreeing on its ease of use [40].

#### Was the Resource Readable?

We also assessed the readability (reading level) of text resources using the Sydney Health Literacy Laboratory (SHeLL) Health Literacy Editor. Optimal readability on the SHeLL Editor is indicated by a score of 8 or below, equating to a grade 8 reading level [41,42]. Generally, health information designed for the general population or patients is recommended to be readable at a grade 8 level or lower [43]. The SHeLL Editor enables the pasting of exact text from a resource into its reading level calculator to provide a specific reading level for text.

#### How Easy Was It to Locate the Resource?

The ease of locating the resource was assessed by determining whether a resource appeared within the first 10 search results on the first page of results on Google. This is based on evidence suggesting the average internet user will only click on results appearing in the first 10 Google search results [36].

#### Was the Resource Evidence-Based?

Resources were evaluated as being evidence-based according to whether or not supporting evidence was cited and accurately represented to support the information they provided, or if they indicated in any background content whether research was involved in the development of the resource content.

#### Was the Resource Developed Through Co-Design With Survivors of AYA Cancer?

Co-design refers to methods used to engage, consult, and work in collaboration with young people to develop research questions, resource content, or interventions [44]. We assessed whether resources were co-designed with survivors of AYA cancer based on whether or not they described using a co-design...
method to develop the information they provide. We note as “unclear” any resources indicating content was developed in consultation with AYAs with cancer, but the exact co-design methods used or extent of engagement with AYAs is not clearly described.

**Was the Resource Culturally Inclusive?**

To our knowledge, no tool exists to assess the cultural inclusivity of international web-based health resources. Therefore, to measure the cultural inclusivity of the resources we identified, we used 7 criteria that were codeveloped by 3 Aboriginal and Torres Strait Islander researchers (including AG) and 2 non-Aboriginal and Torres Strait Islander researchers (including AD) [45]. All researchers involved in the development of these criteria have a strong track record in Indigenous health research [45-51]. Due to the criteria being designed solely for the evaluation of the cultural inclusivity of resources for Aboriginal and Torres Strait Islander peoples in Australia, we only evaluate the cultural inclusivity of resources created by and for Australians [45].

The criteria for cultural inclusivity were as follows [45]: (1) Does the resource include any visual aids (photos, animations, infographics, or charts) that depict or contain information about Aboriginal and Torres Strait Islander peoples? (2) Does the resource include any information or data about Aboriginal and Torres Strait Islander peoples? (3) Does the resource include any Aboriginal and Torres Strait Islander design or artwork? (4) Does the resource provide any evidence of leadership, involvement, or governance by peoples, communities, or organizations that identify as or represent populations that are Aboriginal and Torres Strait Islander? (5) Is the resource available in any Aboriginal and Torres Strait Islander languages? (6) Is any of the language used strengths-based and respectful to Aboriginal and Torres Strait Islander peoples? and (7) Does the resource include a contact (phone number, email, or website) for any culturally relevant or personalized support and information for Aboriginal and Torres Strait Islander peoples?

AG, a Pakana woman from Lutruwita (Tasmania), reviewed all resources to determine their relevancy to Aboriginal and Torres Strait Islander peoples. As this tool is not validated, there is no clear minimum number of criteria that should be achieved to determine the cultural competency of a resource. As such, we report the number of criteria that were met descriptively and describe the strengths and shortfalls of the resources.

**Results**

**Research Question 1: What Web-Based Resources Exist for Engaging With Education or Work After a Cancer Diagnosis, That AYAs With Cancer Are Likely to Encounter When Conducting a Google Search?**

A total of 24 AYA-specific resources met eligibility criteria and were included (Table 2). Most were published by NGOs (n=19, 79%). All resources focused on information provision rather than advertising, support, or intervention, with content shared through text information, text and video stories from other AYAs with cancer, contact details for support organizations, or lists of strategies to navigate education and work challenges. A total of 8 resources were from the United States [52-59], 8 from Australia [60-67], 6 from the United Kingdom [68-73], and 2 from Canada [74,75]. A total of 7 resources focused on education, 8 on work, and 9 on both. There was no difference in search results between searches conducted in Sydney and Perth.

Most resources did not target specific cancer types or stages of the cancer trajectory, although 4 were developed for people diagnosed with blood cancer [59,66,69,74]. There was little consistency in the topics covered across resources, with only a few common topics covered (Table 3).

**Table 2. Cultural inclusivity of resources identified through the environmental scan [45].**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Resources meeting each criterion, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the resource include any visual aids (photos, animations, infographics, or charts) that depict or contain information about Aboriginal or Torres Strait Islander peoples?</td>
<td>0</td>
</tr>
<tr>
<td>Does the resource include any information or data about Aboriginal or Torres Strait Islander peoples?</td>
<td>0</td>
</tr>
<tr>
<td>Does the resource include any Aboriginal or Torres Strait Islander design or artwork?</td>
<td>4</td>
</tr>
<tr>
<td>Does the resource provide any evidence of leadership, involvement, or governance by people, communities, or organizations that identify as or represent populations that are Aboriginal or Torres Strait Islander?</td>
<td>0</td>
</tr>
<tr>
<td>Is the resource available in Aboriginal or Torres Strait Islander languages?</td>
<td>0</td>
</tr>
<tr>
<td>Is any of the language used strengths-based and respectful to Aboriginal and Torres Strait Islander peoples?</td>
<td>6</td>
</tr>
<tr>
<td>Does the resource include a contact (phone number, email, or website) for any culturally relevant or personalized support and information for Aboriginal and Torres Strait Islander peoples?</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 3. Common information and strategies covered in resources.

<table>
<thead>
<tr>
<th>Topic covered</th>
<th>Resource explored topic in relation to education, work, or both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disability rights and accessing accommodations or provisions (n=16) [52-55,58,61-63,65-67,71,72,74]</td>
<td>Both</td>
</tr>
<tr>
<td>Telling people or talking about your cancer (n=13) [52-55,60-63,65-67,70-72]</td>
<td>Both</td>
</tr>
<tr>
<td>Applying for jobs or changing careers (n=11) [52,61,64,66,67,69,71-75]</td>
<td>Both</td>
</tr>
<tr>
<td>Managing physical or cognitive impacts of cancer treatment (n=7) [54,57,65-67,71,75]</td>
<td>Both</td>
</tr>
</tbody>
</table>

**Research Question 2a: Were Resources Understandable?**

Understandability of all text and resources was very good, with all but 1 resource [68] scoring 80% or more on the PEMAT (Table S1 in Multimedia Appendix 1 [52-75]).

**Research Question 2b: Were Resources Actionable?**

Actionability varied greatly across resources, ranging from 40%-100% on the PEMAT (Table S1 in Multimedia Appendix 1 [52-75]). Less actionable text resources tended to be those focused on broad information and strategies, such as tips on how to tell your employer about your diagnosis and the suggestion to seek counseling support through a university campus student services center, rather than advice on when and how AYAs can take specific steps to address their concerns. In general, video resources were the least actionable in that most involved AYA survivors telling their own personal stories related to education and work challenges after a cancer diagnosis, rather than providing advice or strategies to other AYAs.

**Research Question 2c: Were Text Resources Also Readable?**

Readability for all 23 text resources was very poor, with reading levels ranging between grades 8.5 and 16.0 (mean 12, SD 1.97), indicating that, on average, the included resources require completion of a high school degree to comprehend. No resources were assessed to be the optimal reading level of grade 8 or lower (Table S1 in Multimedia Appendix 1 [52-75]).

**Research Question 2d: Were Resources Easy to Locate?**

Resources were difficult to locate, with 22 out of 24 (90%) relevant resources appearing on the second or third page of the Google search results (Table S1 in Multimedia Appendix 1 [52-75]). Results of the Google search prioritized resources related to younger children diagnosed with cancer and their engagement with school, as well as older adults returning to work. AYA-specific resources were scattered in between these less relevant results, as well as other, less relevant results, such as academic journal publications, links to hospital-based cancer services, and information about specific types of cancer.

**Research Question 2e: Were Resources Evidence-Based?**

No resources cited any research-based evidence to support the information provided (Table S1 in Multimedia Appendix 1 [52-75]). Most resources did not describe how the content was developed. Where any description was provided, resources tended to be developed through consultation with expert informants, such as career counselors, oncologists, hematologists, and social workers.

**Research Question 2f: Were Resources Co-Designed?**

No resources specifically discussed co-design methods used to develop resource content in collaboration with AYAs.

**Research Question 2g: Were Resources Culturally Inclusive?**

The number of cultural inclusivity criteria each resource addressed is summarized in Table S1 in Multimedia Appendix 1 [52-75]. The number of resources meeting specific criteria in the cultural inclusivity checklist is summarized in Table 2. Generally, cultural inclusivity of Aboriginal and Torres Strait Islander peoples was very poor, with only 2 out of 7 inclusivity criteria met by 1 or more resources: inclusion of Aboriginal or Torres Strait Islander cultural design or artwork and an acknowledgment on the web page recognizing Aboriginal and Torres Strait Islander peoples (minimally reflecting the use of strengths-based language that is respectful of Aboriginal and Torres Strait Islander peoples).

In reviewing all resources, both Australian and internationally-designed, we noticed several further equity, access, diversity, and representative issues with resources that were not initially part of our aims, that are important to highlight. Pictures and videos presented in resources almost exclusively portrayed heterosexual relationships, women were more commonly represented in pictures than men, and women were more commonly shown to be young, White, and of thinner build. Settings also showed middle-class, suburban, or urban areas rather than lower-socioeconomic, rural, or remote settings. Most resources also primarily assumed internet access and support from family or friends were available to AYAs. Lastly, language was not gender neutral and tended to assume heterosexual, 2-parent families.

**Discussion**

**Overview**

This environmental scan aimed to (1) determine what web-based resources exist to support survivors of AYA cancer in their engagement with education or work and (2) assess the understandability, readability, actionability, and cultural inclusivity of resources, as well as how easy resources were to locate, whether they provide evidence-based information, and whether they were co-designed with survivors of AYA cancer. We found few high-quality resources on the topic of returning to education or working for AYAs with cancer in Australia.
Although the understandability of most resources was high, the readability of text-based resources was poor, with most text resources requiring reading levels at the university education level or higher. This discrepancy may be due to the understandability criteria being quite broad (eg, text “material uses common, everyday language” or “material ‘chunks’ information into short sections”) and not directly providing criteria against which age-appropriate understandability could be assessed. For example, a resource might include everyday language for a young adult in small sections, but sentence structures or legal terminology related to education or work rights may be more complicated, thus affecting readability. Our findings on poor readability of resources are consistent with literature indicating most AYAs with cancer who access web-based resources report the resources require high health literacy and present information that is difficult to understand, critically evaluate, and act on [76,77].

We also found most web-based resources limited in their modes of information provision, primarily using text to provide lists of information and strategies. Few resources involved audio-visual content that may be preferable for the AYA population [35]. Where audio-visual content was provided, it was often focused on individual stories and experiences rather than the provision of guidance to AYAs and actionable strategies to navigate the return to education or work after cancer.

It is therefore unsurprising that the actionability of resources was moderate, with resources scoring 60%-100% on actionability. The focus on broad information and strategies in most of the resources reviewed may feel overwhelming to AYAs, given that it can be difficult for a young person to review a long list of suggestions and determine what they should act on given their individual health status, needs, and education or work goals [25]. Previous studies have provided evidence that Australian survivors of AYA cancer report low confidence in their ability to assess the reliability and validity of health-related information [77]. From a developmental standpoint, adolescents may not have fully developed their critical thinking skills yet [77]. This underscores the importance of providing information to AYAs that is both understandable and actionable.

It is also important to note that no resources cited an evidence base (ie, peer-reviewed scientific literature) for their information or recommendations, nor did any resource specify co-design of any content with survivors of AYA cancer. Instead, most resources assumed a certain level of understanding and self-motivation to act on the provided information or strategies. Most resources were also difficult to locate, appearing on multiple pages in a Google search. However, reliance on evidence-based information and co-design of resources with AYA survivors are widely acknowledged as critical to ensuring that information content and delivery methods are optimized for the specific needs of this age group [78-81].

Lastly, few resources met more than 1 criterion for cultural inclusivity. The lack of culturally inclusive resources for Aboriginal and Torres Strait Islander peoples may exacerbate existing health inequalities in people with cancer [82]. While there is no research, to our knowledge, describing specific concerns related to return to work or education for Aboriginal and Torres Strait Islander AYAs with cancer, financial distress is a common area of unmet need for Aboriginal and Torres Strait Islander adults with cancer in Queensland [83]. Considering the employment and educational disparities that are known within this population broadly [84,85], it stands to reason that Aboriginal and Torres Strait Islander AYAs with cancer may be in particular need of information to support their educational and employment endeavors. However, resources were generally not inclusive of Aboriginal and Torres Strait Islander peoples.

In turn, the literature suggests that such limited inclusivity in health information can lead to feelings of isolation, feeling misunderstood by health services, and reduced self-efficacy in patients to follow medical advice [86]. There is an urgent need to address this gap in resources available to Aboriginal and Torres Strait Islander survivors of AYA cancer.

### Comparison to Previous Work

While some NGOs and health care institutions provide lists of web-based resources for AYAs with cancer on their website, no previous research has specifically collated and evaluated web-based resources to support AYAs’ return to education or work after cancer. This environmental scan is the first to evaluate the age appropriateness, accessibility, understandability, and cultural inclusivity of web-based resources specifically targeting the education and work needs of AYAs with cancer in Australia.

### Limitations

There were some limitations worth noting. The search strategies used were constructed to mirror typical searches AYAs might conduct with a select set of keywords. However, these strategies may not capture all modes of searches AYAs might conduct, such as asking questions in the Google search or using other search terms not featured, and we did not include results such as blogs or social media posts from which AYAs may also seek information. We only conducted searches in English and did not find or include resources published in other languages that may be relevant. Furthermore, there are some limitations associated with assessing the cultural inclusivity of international resources using a tool designed specifically for Australia. However, it is important to note that, to our knowledge, no international tools exist to assess the cultural inclusivity of international resources for Aboriginal and Torres Strait Islander AYAs. We also aimed to optimize the reach of our searches by using 24 unique search term combinations, as well as by conducting the searches from both Sydney (where the investigators are located) and Perth (through a virtual private network). Finally, given that AYAs use web-based resources and information, an important next step will be understanding what AYAs themselves think about available web-based resources in terms of appropriateness or usefulness, which was beyond the scope of the current environmental scan conducted here.
Conclusions

AYAs diagnosed with cancer frequently turn to the internet to seek information related to their diagnosis, treatment, and psychosocial needs [24,25,76,87]. Information accessed on the internet can play a major role in AYAs’ decisions to seek care or support to address their specific needs and concerns [24,25,76,87]. Findings from this environmental scan suggest AYAs diagnosed with cancer in Australia would benefit from more tailored, evidence-based, and culturally inclusive web-based resources that are easy to locate, are provided in multiple formats (eg, text as well as audio-visual), are presented at the reading level of someone in year 8 or below and are easy to act on. While some resources describe their development as being done in consultation with survivors of AYA cancer, it is unclear to what extent a co-design approach was taken. A co-design approach would be beneficial to at least ensure the understandability, readability, actionability, and cultural inclusivity of any future resources developed.

Acknowledgments

This paper was prepared with funding support from the Cancer Australia Supporting People With Cancer grant (CA-ITA-2021/01) and the Kids Cancer Project and Canteen. JEF is a Maridulu Budyari Gumal (SPHERE) Cancer CAG Senior Research Fellow and is supported by a cancer institute NSW Research Capacity Building grant (2021/CBG003). AD is funded by a University of Queensland Faculty of Medicine Fellowship. UMS-D is supported by an Early Career Fellowship from the Cancer Institute of New South Wales (2020/ECF1163). The Behavioral Sciences Unit is proudly supported by the Kids with Cancer Foundation, by the Kids Cancer Alliance, as well as a Cancer Council New South Wales Program grant (PG16-02) with the support of the estate of the late Harry McPaul.

Data Availability

The data sets generated during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors contributed to project conceptualization, methodology, writing, reviewing, and editing. Authors CES, JEF, GD, AD, and AG contributed to data curation and formal analysis. CES wrote the original manuscript with supervision from JEF.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Resources identified. 

References


57. College students living with a chronic condition. Center for Online Education. URL: https://www.fors-students/chronic-health-issues/ [accessed 2024-02-21]
68. What should be considered when planning to return to work after cancer? Shine Cancer Support. URL: https://www.youtube.com/watch?v=f5_41vEthOE [accessed 2024-02-21]
74. Returning to work or school after cancer treatment. Leukemia and Lymphoma Society. URL: https://www.youtube.com/watch?v=BHVlZeZotX8 [accessed 2024-02-21]
75. The POGO school and work transitions program. Pediatric Oncology Group of Ontario. URL: https://www.pogo.ca/programs-support/survivor-care/transitions-school-work-counselling/ [accessed 2024-02-21]


Abbreviations

AYA: adolescent and young adult
NGO: nongovernmental organization
PEMAT: Patient Education Materials Assessment Tool
SHeLL: Sydney Health Literacy Laboratory
Evaluation of the e–Mental Health Intervention Make It Training From Patients' Perspectives: Qualitative Analysis Within the Reduct Trial

Julia Barbara Krakowczyk1,2,3, MSc; Femke Truijens4, PhD; Martin Teufel1,2,3, MD; Tania Lalgi1,2,3, BSc; Jana Heinen5,6, MSc; Caterina Schug7,8, MSc; Yesim Erim7,8, MD; Michael Pantförder9, MSc; Johanna Graf5,6, PhD; Alexander Bäuerle1,2,3, PhD

1Clinic for Psychosomatic Medicine and Psychotherapy, LVR-University Hospital Essen, University of Duisburg-Essen, Essen, Germany
2Comprehensive Cancer Center, University Hospital Essen, Essen, Germany
3Center for Translational Neuro- and Behavioral Sciences, University of Duisburg-Essen, Essen, Germany
4Department of Psychology, Educational and Child Studies, Erasmus University Rotterdam, Rotterdam, Netherlands
5Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tübingen, Eberhard Karls University Tübingen, Tübingen, Germany
6Comprehensive Cancer Center, University Hospital Tübingen, Tübingen, Germany
7Department of Psychosomatic Medicine and Psychotherapy, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany
8Comprehensive Cancer Center, University Hospital Erlangen, Erlangen, Germany
9Fraunhofer Institute for Software and Systems Engineering, Dortmund, Germany

Corresponding Author:
Julia Barbara Krakowczyk, MSc
Clinic for Psychosomatic Medicine and Psychotherapy
LVR-University Hospital Essen
University of Duisburg-Essen
Virchowstraße 174
Essen
Germany
Phone: 49 201438755216
Email: julia.krakowczyk@uni-due.de

Abstract

Background: Make It Training is an e–mental health intervention designed for individuals with cancer that aims to reduce psychological distress and improve disease-related coping and quality of life.

Objective: This study evaluated the experienced usefulness and usability of the web-based Make It Training intervention using a qualitative approach.

Methods: In this study, semistructured interviews were conducted with participants at different cancer stages and with different cancer entities. All participants had previously taken part in the Reduct trial, a randomized controlled trial that assessed the efficacy of the Make It Training intervention. The data were coded deductively by 2 independent researchers and analyzed iteratively using thematic codebook analysis.

Results: Analysis of experienced usefulness resulted in 4 themes (developing coping strategies to reduce psychological distress, improvement in quality of life, Make It Training vs traditional psychotherapy, and integration into daily life) with 11 subthemes. Analysis of experienced usability resulted in 3 themes (efficiency and accessibility, user-friendliness, and recommendations to design the Make It Training intervention to be more appealing) with 6 subthemes. Make It Training was evaluated as a user-friendly intervention helpful for developing functional coping strategies to reduce psychological distress and improve quality of life. The consensus regarding Make It Training was that it was described as a daily companion that integrates well into daily life and that it has the potential to be routinely implemented within oncological health care either as a stand-alone intervention or in addition to psychotherapy.

Conclusions: e–Mental health interventions such as Make It Training can target both the prevention of mental health issues and health promotion. Moreover, they offer a cost-efficient and low-threshold option to receive psycho-oncological support.
background
Cancer is one of the leading causes of death worldwide, and its prevalence is constantly increasing [1]. Worldwide, 19.3 million new cases of cancer were diagnosed in 2020 [2]. By 2024, a total of 27.5 million new cases of cancer are expected each year [2]. Receiving a cancer diagnosis and undergoing cancer treatment are associated with a high psychological burden [3,4]. Approximately every second individual diagnosed with cancer experiences high psychological distress, and one-third of all individuals across different cancer stages and types meet the criteria for at least one mental health disorder [5–7].

Due to the high psychological burden associated with cancer, a significant number of individuals seek psycho-oncological support [8–10]. Previous research has proven the efficacy of psycho-oncological treatment on different outcomes such as distress, fatigue, depression, anxiety, and quality of life [11–16]. However, receiving proper psycho-oncological support is difficult due to various barriers within the health care system [10,17,18]. These include geographic barriers, the stigma of seeking mental health services, financial constraints, continuity of health care, and the limited availability of mental health professionals [19–21]. Thus, efforts are required to expand access to mental health support for patients with cancer [4,8].

eHealth interventions offer a cost-efficient approach to overcome barriers in psycho-oncological care [16,22,23]. Most of these eHealth interventions consist of (web) applications that are based on psychotherapeutic approaches such as cognitive behavioral therapy (CBT) [24–28]. Existing research has demonstrated the efficacy of psychological eHealth interventions for individuals with cancer on outcomes such as distress, depression, anxiety, fatigue, and quality of life [16,25–27].

Most of the studies evaluating psycho-oncological eHealth interventions have proven their efficacy by adapting a quantitative research approach [16,25–27], wherein statistical analyses are conducted to investigate the pre- and postintervention scores of standardized questionnaires to assess statistically significant differences [29]. Although this approach is considered the gold standard for efficacy research, it does have some limitations [30]. These limitations include missing information on individual experiences, as well as missing in-depth information on the mechanisms behind the change that led to the statistical significance displayed in the data [31]. The inclusion of qualitative research offers an in-depth understanding of the mechanisms [32–34]. Considering research findings from both qualitative and quantitative approaches allows for a more holistic understanding of not only whether an intervention works but also how and why [35,36]. Thus, it offers in-depth knowledge of change mechanisms and the possibility of optimizing existing interventions. Moreover, assessment of eHealth interventions using a mixed methods approach is associated with increased adaptation to patients’ needs and demands compared to solely using quantitative assessments [37–39].

This paper reports qualitative analyses conducted as part of the Reduct trial (German Clinical Trial Register DRKS00025213) [40]. The Reduct trial is a multicenter randomized controlled trial to assess the efficacy of the web-based Make It Training intervention (mindfulness and skill-based distress reduction training in oncology). To date, it is one of the largest efficacy trials in the field of psycho-oncology. Make It Training is a self-guided (web-based) application aimed at reducing distress in individuals with cancer [40,41]. It is based on CBT, acceptance and commitment therapy (ACT), and mindfulness-based stress reduction (MBSR). Over 4 months, individuals are supported by Make It Training through skill training, psychoeducation, interactive exercises, mindfulness, and psychotherapeutic techniques. Make It Training aims to reduce psychological distress, improve disease-related coping, and improve quality of life. It was developed to bridge the gap in the lack of psycho-oncological support in the health care system that currently exists in certain regions. The papers by Bäuerle et al [40] and Heinen et al [41] outline the study and intervention protocols, respectively.

study objectives
Taking on a qualitative stance, this study examined the experienced usefulness and usability of Make It Training from patients’ perspectives. The aim of this study was to obtain a more holistic view and enrich the understanding of individuals’ experiences concerning Make It Training beyond the boundaries of quantitative data [35,36]. When referring to the experience of usefulness, this study took on a psychotherapeutic perspective and referred to the patients’ general evaluation of Make It Training, changes experienced while completing the intervention, attribution of these changes, specific aspects of the intervention that they found particularly useful or hindering, and recommendation to other individuals with cancer. On the basis of the study by Gould and Lewis [42] and the Health IT Usability Evaluation Model [43], the term usability comprises the patients’ experienced user-friendliness, efficiency, accessibility, and practicability of the intervention.

methods
study design and procedure
This study was based on the guidelines of Levitt et al [44] and the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [45]. It consisted of one-on-one semistructured interviews. The interviews were conducted by a trained female interviewer who was experienced with qualitative research. To avoid any potential bias, the interviewer was not part of the core research team of the Reduct trial. There was no previous relationship established between the interviewer and the participants before the study began. Moreover, the
participants did not have personal knowledge of the researcher. In total, 33% (2/6) of the participants completed the interviews in person, and 67% (4/6) did so digitally. Apart from the interviewer and the interviewee, there was no other person present during the interviews. All participants were interviewed once. To focus on the dialogue between the interviewee and the interviewer, no field notes were taken during the interviews. No transcripts were returned to the participants for comments or corrections. The COREQ checklist can be found in Multimedia Appendix 1 [45].

Recruitment
The participants of 1 study center that completed the Make It Training intervention within the Reduct trial [40] between May 2022 and September 2022 were contacted via email and telephone and invited to participate in this study. Purposive sampling (ie, completion of Make It Training) was carried out to obtain information-rich participants as well as in-depth experiences with Make It Training [34,46,47]. Recruitment took place in an early phase of the Reduct trial, so 11 participants were eligible to be contacted in total. Of these 11 participants, 5 (45%) either did not respond or could not participate for personal reasons. The final sample consisted of 6 participants. On the basis of Crouch and McKenzie [48], a small sample size was selected to put emphasis on the relationship between the researcher and the participant, as well as to explore the patients’ lived experiences with Make It Training in depth.

For the inclusion, exclusion, and completion criteria (eg, current cancer diagnosis, command of the German language, internet connection, age of >18 years, and no psychotherapy during the intervention period) of the Reduct trial, we refer to the study protocol by Bäuerle et al [40]. This study was based on the inclusion and exclusion criteria of the Reduct trial.

Ethical Considerations
This study was approved by the Ethics Committee of the Medical Faculty of the University of Duisburg-Essen (22-10,902-BO). All interviews were conducted on the premises of the university and audiotaped with the interviewees’ consent. The data were pseudonymized. The data protection–compliant audio files and identifying information were stored in a password-protected database. After providing written informed consent, the participants were interviewed. The participants had the option to be interviewed either in person at the clinic or digitally through a data protection–compliant software for clinicians [49]. There was no compensation or any form of reimbursement.

Semistructured Interview
The interview questions were divided into 9 segments. The first segment focused on explaining the study background and gathering sociodemographic information. In the second to ninth segments, interviewees were asked about the following: general experience with Make It Training, changes that they noticed since completing the intervention, attribution of these changes, content of the intervention that they perceived as particularly helpful or not helpful, content that was perceived as missing, the motivation to participate in the intervention, usability, and recommendation of the intervention to other individuals with cancer.

The interview questions were developed based on the Client Change Interview (CCI) [50] and the Health IT Usability Evaluation Scale (Health ITUES) [51]. The CCI was chosen as it is an established interview within psychotherapy research to assess self-perceived changes and attribution of changes related to psychotherapy [50]. In addition, it helps to identify perceived helpful or unhelpful components of psychotherapeutic interventions [50].

The Health ITUES is a questionnaire used to evaluate the usability of eHealth technologies among people with chronic diseases [52]. It was chosen as it is a validated assessment instrument to evaluate the feasibility and usability of eHealth interventions.

The full version of the semistructured interview is provided in Table S1 in Multimedia Appendix 2 [50,51]. In addition, self-generated questions were included (Textbox 1).

Textbox 1. Self-generated questions of the semistructured interview.

- How did you perceive the operation and user-friendliness of the Make It Training?
- How did you perceive the additional service in the form of reminder emails and contacts in the event of technical difficulties?

Data Analysis
The data were analyzed using thematic codebook analysis [53,54]. Thematic analysis was chosen due to its wide application across paradigms [54-56]. An overall deductive approach was chosen because it is an established approach to evaluate user experiences with digital interventions [57]. Moreover, it is helpful in organizing and categorizing meaningful data in conjunction with the existing literature [34,35,54]. The data were coded partly deductively by 2 independent researchers in 2 rounds of analysis. As the research team was interested in the participants’ in-depth lived experiences with Make It Training rather than general thematic cohesion over the sample, a bottom-up inductive analysis was conducted first, which was then captured in the deductive structure in the second round of analysis.

The analyses were conducted iteratively; that is, they were carried out in a cyclical manner to refine and deepen the understanding of the data through the following steps:

1. Each coder open coded the first 2 transcripts, and individual memos were written.
2. The codes were compared and revised through multiple iterative rounds among the research team to obtain different perspectives. Both coders met to compare their findings, particularly regarding the codes; discuss discrepancies to
ensure consensus on the application of finalized codes and, if applicable, add new codes; and develop a codebook.

3. Both coders agreed that saturation had been attained in the first 2 open-coded transcripts.

4. The finalized codes were divided into categories and themes and tested on the 4 remaining transcripts.

Chronemics (such as hesitation or silence) were taken into account as nonverbal information in the analysis. Overall, there was a high level of agreement (approximately 70%) between the researchers during the evaluation process, and discrepancies were critically discussed during meetings with the research team to reach a consensus. For publishing purposes, all interview quotes were translated from German into English, and the analysis process was reviewed by the research team. All interviews were transcribed using the f4x transcription software and then analyzed using the MAXQDA computer program (VERBI GmbH) [58]. On the basis of the decision to include a small sample size, the research team defined saturation according to Legard et al [59], meaning that saturation was assessed based on whether there was a consensus among the participants regarding the general evaluation of Make It Training and whether the research team felt that they had reached an understanding of the participants’ lived experiences with Make It Training.

Quality Control
All researchers involved had a background in clinical psychology, psycho-oncology, psychosomatic medicine, and psychotherapy with different research experiences (full-time professors, assistant professors, postdoctoral researchers, PhD candidates, and graduate students).

On the basis of Creswell and Miller [60], validity guidelines were followed to ensure the validity of this study. These included triangulation by searching for convergence among diverse sources of information (eg, the lens of the researcher and systematic paradigm) to form themes or categories in a study [60]. Finally, validation procedures included seeking assistance through peer debriefing, which was realized by involving an auditor. The auditor was a senior qualitative researcher with extensive experience in clinical psychology and efficacy research but without familiarity with the Reduct trial and the Make It Training intervention. They audited the first round of findings by reading written findings, questioning the researchers on their procedures, and challenging interpretations and thematic structure. Subsequently, the researchers conducted another iterative round of analysis to synthesize and sensitize the data and fine-tune the findings accordingly. To establish credibility, we ensured to provide a thick and rich description of the setting, participants, and themes of the qualitative study [61].

Results
Overview
A total of 6 (mean 34 min, SD 7 min 56 s; range 20-45 min) one-on-one interviews were conducted. The demographic characteristics of the participants are presented in Table 1.
Table 1. Sociodemographic and diagnosis-related characteristics of the participants (N=6).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Identified as a woman</td>
<td>4 (67)</td>
</tr>
<tr>
<td>Identified as a man</td>
<td>2 (33)</td>
</tr>
<tr>
<td><strong>Age range (y)</strong></td>
<td></td>
</tr>
<tr>
<td>49-56</td>
<td>4 (67)</td>
</tr>
<tr>
<td>57-66</td>
<td>2 (33)</td>
</tr>
<tr>
<td><strong>Cancer type</strong></td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>1 (17)</td>
</tr>
<tr>
<td>Lymphatic; blood-forming tissue</td>
<td>1 (17)</td>
</tr>
<tr>
<td>Skin cancer</td>
<td>2 (33)</td>
</tr>
<tr>
<td>Colon cancer</td>
<td>1 (17)</td>
</tr>
<tr>
<td>Musculoskeletal tumors</td>
<td>1 (17)</td>
</tr>
<tr>
<td><strong>Year of initial cancer diagnosis</strong></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>1 (17)</td>
</tr>
<tr>
<td>2018</td>
<td>1 (17)</td>
</tr>
<tr>
<td>2019</td>
<td>1 (17)</td>
</tr>
<tr>
<td>2020</td>
<td>2 (33)</td>
</tr>
<tr>
<td>2022</td>
<td>1 (17)</td>
</tr>
<tr>
<td><strong>Recurrence</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5 (83)</td>
</tr>
<tr>
<td>No</td>
<td>1 (17)</td>
</tr>
<tr>
<td><strong>Metastasis</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3 (50)</td>
</tr>
<tr>
<td>No</td>
<td>3 (50)</td>
</tr>
</tbody>
</table>

\( ^a \)Sociodemographic characteristic.  
\( ^b \)Medical characteristic and etiopathology.

Theme Classification

Overview

The previously selected categories were divided into 7 themes that were used to focus the qualitative analyses. The themes were used deductively to select excerpts in the interviews that appeared relevant to these themes. Within the selections per theme, excerpts were coded using line-by-line coding and grouped to form information-rich subthemes. All themes and subthemes are reported in the following sections using representative quotes. Further information on the theme classification can be found in Figure 1, whereas Table S2 in Multimedia Appendix 2 summarizes all representative quotes.

The consensus regarding the Make It Training intervention was that it was described as a “daily companion” that integrates well into daily life and that it has the potential to be routinely implemented within oncological health care either as an intervention itself or in addition to psychotherapy (Table S2 in Multimedia Appendix 2, quote 1).
Category 1: Experienced Usefulness

Theme 1: Developing Coping Strategies to Reduce Psychological Distress

Overview Theme 1
This theme is centered on the development of functional coping strategies that participants described as a change related to Make It Training. All participants reported that Make It Training helped them develop a repertoire of coping strategies, which was helpful in reducing psychological distress.

For example, the improvement in emotion regulation was described as such a strategy (Table S2 in Multimedia Appendix 2, quote 2). Another commonly described coping strategy was redefining the relationship with cancer (Table S2 in Multimedia Appendix 2, quote 3).

Subtheme 1.1: Mindfulness Exercises
The increased practice of mindful behavior stood out as a described coping strategy, and it was attributed to the mindfulness exercises provided in Make It Training. The participants strongly embraced the variety of mindfulness exercises provided in the intervention. Interviewee 6 would have preferred even more exercises within Make It Training (Table S2 in Multimedia Appendix 2, quote 4).

The mindful breathing exercises were most commonly described as helpful. They were perceived as a new coping skill that could be integrated into daily life for stress management and tension reduction. One of the participants also positively noted the long-term advantages of breathing exercises (Table S2 in Multimedia Appendix 2, quote 5). This statement illustrates the advantages of mindful breathing exercises as part of the coping repertoire. Moreover, it demonstrates the practical application of the techniques in daily life as well as the interviewees’ subjective perception of improvement.

Subtheme 1.2: Initiating Introspection
Most participants reported that Make It Training helped initiate introspection, which was described as supportive in dealing with difficult situations. It was further described as developing the skill to observe and interpret one’s own thinking patterns, emotions, and behavior and not just be overwhelmed by them (Table S2 in Multimedia Appendix 2, quotes 6 and 7). Moreover, being able to observe one’s inner world (ie, introspection) can help shift attention to positive aspects in difficult phases (Table S2 in Multimedia Appendix 2, quote 8).

Subtheme 1.3: Psychoeducation Increased Understanding of Psychological Distress Associated With Cancer and Communication About the Illness
Many participants experienced the psychoeducational components within the intervention as helpful because they led to a better understanding of cancer and its associated psychological distress and somatic restrictions. The participants reported that they were able to learn not only about personal circumstances but also how to communicate better and more effectively approach family members. In this regard, the expert videos provided, where health care professionals reported on each topic, were perceived as useful (Table S2 in Multimedia Appendix 2, quote 9).

Theme 2: Improvement in Quality of Life

Overview Theme 2
All participants reported that Make It Training helped increase their quality of life. This was described as redefining...
perspectives on life circumstances and cancer. Moreover, health-related behavior change, increase in resilience, and enhanced practice of mindful behavior were described as positively contributing to quality of life (Table S2 in Multimedia Appendix 2, quote 10).

**Subtheme 2.1: Cognitive Restructuring and Changing Perspective on Life With Cancer**

Participants reported that Make It Training helped modulate existing thinking patterns. This was commonly described as changing perspectives on life with cancer, as well as on the cancer diagnosis itself (Table S2 in Multimedia Appendix 2, quote 11). Another participant described a redefined relationship with pain (Table S2 in Multimedia Appendix 2, quote 12).

**Subtheme 2.2: Building Resilience**

Participants reported that Make It Training helped them become more resilient, which was described as developing the ability to better deal with unpleasant situations such as chemotherapy (Table S2 in Multimedia Appendix 2, quote 13).

**Subtheme 2.3: Initiating a More Relaxed State in Daily Life**

The participants described that the intervention was helpful to experience a more relaxed state in daily life, which positively contributed to their quality of life (Table S2 in Multimedia Appendix 2, quotes 14 and 15).

**Theme 3: Make It Training Versus Traditional Psychotherapy**

**Overview Theme 3**

While evaluating the Make It Training intervention, some participants drew a comparison between Make It Training and traditional psychotherapy. In total, 33% (2/6) of the participants had previous psychotherapeutic experience. Even though Make It Training was perceived as a helpful and easily accessible format to receive psycho- oncological support, 83% (5/6) of the patients reported that it did not replace traditional psychotherapy (Table S2 in Multimedia Appendix 2, quote 16). In contrast, one participant reported preferring Make It Training to traditional face-to-face psychotherapy (Table S2 in Multimedia Appendix 2, quote 17).

**Subtheme 3.1: Recommendation to Other Patients**

All participants had been diagnosed with different cancer entities and stages (Table 1). Overall, all reported recommending Make It Training to others as they were convinced that other individuals with cancer could benefit from the intervention as well. Some of them suggested that a psycho-oncological eHealth intervention such as Make It Training should be offered as a routine intervention within oncological health care.

Multiple participants argued that particularly individuals with a first-time cancer diagnosis would substantially benefit from the intervention. One participant hypothesized that providing individuals with a first-time diagnosis of cancer with an eHealth application such as Make It Training would help them process and better deal with the cancer diagnosis (Table S2 in Multimedia Appendix 2, quotes 18 and 19).

**Subtheme 3.2: Communication With Therapist and Request for Blended Therapy Format**

Make It Training is a purely self-guided eHealth intervention. Some participants wished for more communication with a therapist. In this context, they stressed the importance of a patient-therapist interaction. Some participants reported that Make It Training might be even more beneficial with additional therapist guidance. In this regard, additional therapist consultations via phone or email were suggested. Moreover, participants reported that these options would offer the opportunity to better voice challenges, misunderstandings, and questions. A total of 50% (3/6) of the participants expressed a preference for a blended format (ie, a combination of Make It Training with traditional face-to-face psychotherapy; Table S2 in Multimedia Appendix 2, quote 20).

**Theme 4: Integration Into Daily Life**

**Overview Theme 4**

The intervention was described as a “daily companion” (interviewee 4) or “a wonderful companion for everyday life” (interviewee 2) that could help a lot of individuals with cancer. Make It Training provided participants with a variety of psychoeducational information, psychotherapeutic exercises, and skill training that were perceived as suitable for integration into daily life. All participants reported that they had incorporated the received information or skills that they found valuable and implementable (see Table S2 in Multimedia Appendix 2, quotes 21 and 22, for examples of how participants integrated the skills into their daily lives).

**Subtheme 4.1: Motivation**

In the initial phase, all participants reported being motivated to complete the intervention. However, there were divided opinions regarding motivation after that initial phase. Some experienced Make It Training to be action activating because “it was a meaningful engagement with the disease” (interviewee 2). For others, the motivation gradually declined.

One participant brought up an analogy from sports to describe their motivation. They addressed the fact that, over time, they lacked the motivation to continue through Make It Training. However, the reminder emails helped keep the participant motivated (Table S2 in Multimedia Appendix 2, quote 23). In contrast, there were participants who did not need an external motivator (Table S2 in Multimedia Appendix 2, quotes 24-26).

**Subtheme 4.2: Difficulty Level of Yoga Exercises**

Make It Training comprised physical exercises in the form of yoga. There were mixed opinions on the difficulty level of these exercises as some participants perceived them as physically exhausting, whereas others did not. An older participant reported that some physical exercises were too straining due to restrictions caused by a lack of mobility because of the cancer (Table S2 in Multimedia Appendix 2, quote 27).

**Subtheme 4.3: High Curiosity When Completing the Make It Training Intervention**

Curiosity was high among all participants to see “what’s new there?” (interviewee 1) when a new module was unlocked. Curiosity was described as high because one had to wait a week...
to unlock a new module, which was perceived as exciting (Table S2 in Multimedia Appendix 2, quotes 28-30). Overall, participants seemed to support the format in which content is unlocked incrementally as it generates curiosity.

Category 2: Usability

Theme 5: Efficiency and Accessibility of the Make It Training Intervention

Overview Theme 5
The digital setup allowed all participants to work through the modules independent of time and place. Because of that, Make It Training was perceived as an efficient and easily accessible format to receive psycho-oncological support (Table S2 in Multimedia Appendix 2, quote 31).

Subtheme 5.1: Low-Threshold and Trustworthy Accessibility of Psychological Support
The content provided during the intervention was perceived as professional and trustworthy. It was reported that having access to Make It Training was not associated with barriers that were previously experienced by some participants when seeking psychotherapy. This was perceived as very positive (Table S2 in Multimedia Appendix 2, quote 31).

Subtheme 5.2: Retrievability of Content Independent of Time and Place
All participants positively outlined the retrievability of the content. This refers to the possibility to flexibly retrieve the contents of Make It Training independent of time and place. When a module is activated, the participants can choose when and for how long they want to work on it, as well as on what parts. This was perceived as useful as it offers the flexibility to work on the modules independently of physicians’ appointments, operations, or other medical examinations. Thus, Make It Training was considered “really timely-ideal” (interviewee 4; Table S2 in Multimedia Appendix 2, quote 35).

Participants also reported that the retrievability of the content helped them assess whether a skill that was learned could actually be internalized as well, which was perceived as a benefit (Table S2 in Multimedia Appendix 2, quotes 36 and 37).

Theme 6: User-Friendliness

Overview Theme 6
There were mixed opinions regarding the user-friendliness of Make It Training. Overall, participants considered the application user-friendly. One of the most common reasons why the intervention was described as user-friendly was that it was perceived as not requiring much guidance when using it (Table S2 in Multimedia Appendix 2, quote 38).

One participant criticized the user-friendliness of Make It Training because they perceived the software interface as confusing (Table S2 in Multimedia Appendix 2, quote 39).

Subtheme 6.1: Customization of the Modules
Make It Training follows a certain chronology in the order of the modules, which is not customizable. This was experienced by most participants as very limiting, and they would have liked to be able to work through the modules in their own order (Table S2 in Multimedia Appendix 2, quote 40).

Subtheme 6.2: Software Interface
There were mixed opinions regarding the software interface of Make It Training. Some participants perceived the layout of Make It Training as clear and stimulating. In contrast, others pointed out the unclear and childish presentation of the modules. One participant also came up with an analogy to a “kids board game” (interviewee 5). In general, the rather playful approach was appreciated (Table S2 in Multimedia Appendix 2, quotes 41 and 42).

Subtheme 6.3: Email Reminder to Increase Adherence
There were mixed opinions regarding the reminder emails that all participants received throughout the intervention. Most perceived them as a helpful addition that encouraged them; however, some of the participants perceived them as a bother (Table S2 in Multimedia Appendix 2, quote 43).

Subtheme 6.4: Technical Aspects
Most of the participants did not report any significant technical difficulties or perceived deficiencies. Common technical issues included internet connection or low-resolution quality of the videos.

Theme 7: Recommendations to Design the Make It Training Intervention to Be More Appealing
The participants gave feedback on how to design the Make It Training intervention to be more appealing. One module that focused on the family members of individuals with cancer was regarded by 33% (2/6) of the participants as lacking sensitivity. They reported that working through this module seemed inappropriate and upsetting for those without family members (Table S2 in Multimedia Appendix 2, quotes 44 and 45).

As another recommendation, some participants expressed the need to adapt the modules to the stage of cancer and the current treatment phase (Table S2 in Multimedia Appendix 2, quote 46).

Regarding usability, participants reported minor technical issues or design shortcomings that affected their navigation of and interaction with the program (eg, struggle to remember their position or progress within the program and challenges in finding the right areas to click or interact with). Clearer indicators or visual cues to help users track their progress and easily identify their current location within the program’s content or structure were suggested (Table S2 in Multimedia Appendix 2, quotes 47-49).

Discussion

Principal Findings
This study examined the experienced usefulness and usability of Make It Training from patients’ perspectives using a qualitative approach, which was accomplished through thematic analysis of interviews conducted with individuals with cancer at different stages of severity. Analysis of their experience of the usefulness of Make It Training resulted in 4 themes (developing coping strategies to reduce psychological distress,
improvement in quality of life, Make It Training vs traditional psychotherapy, and integration into daily life) with 11 subthemes. Analysis of their experienced usability resulted in 3 themes (efficiency and accessibility, user-friendliness, and recommendations to design the Make It Training intervention to be more appealing) with 6 subthemes. All participants positively evaluated Make It Training. Moreover, all participants reported that they experienced positive changes while completing the Make It Training intervention and attributed these changes to the intervention itself. The overall usability of Make It Training was experienced as positive as well, although the experiences showed variation due to personal preferences. Overall, the results of this study point to a high satisfaction with Make It Training.

The themes that were discussed as perceived changes during the Make It Training intervention are consistent with its overall goal, which is to support individuals with cancer with disease-related coping, improvement in quality of life, and reduction in psychological distress [40,41]. Moreover, the aforementioned results are in line with those of the study by Ringwald et al [62], who assessed the acceptance of and satisfaction with a previous version of the Make It Training intervention in a pilot study. In this study, the acceptance and satisfaction rates of Make It Training were high, and 87% of the participants reported that they would recommend the intervention to other individuals with cancer [62]. Overall, the results from both the study by Ringwald et al [62] and our study point to a high acceptance of and satisfaction with Make It Training. Because of their satisfaction with Make It Training, the participants stated that it should be implemented as a routine intervention within health care. Previous research has shown that there is a relationship between acceptance of eHealth interventions and their actual use [63-67]. Acceptance is also an important factor for adherence [68]. Thus, given the acceptance of and satisfaction with Make It Training, it might have potential as an eHealth intervention to be routinely implemented in oncological health care as a medical device. In Germany, for example, there is a more recent regulation that eHealth interventions can be prescribed by health care professionals.

The Make It Training was described as a low-threshold and efficient format to receive psycho-oncological support. This was perceived as extraordinarily helpful as some participants had previously experienced difficulties with receiving proper psycho-oncological support, which is known to be a common problem in certain regions [10,17,18]. In this regard, the retrieval of content independent of time and place was described as being helpful with internalizing learned skills and accessing psychological support quickly when needed. These results further support the implementation of eHealth interventions such as Make It Training as an integral part of oncological health care. Digital interventions, if they are accepted among users, can overcome barriers associated with receiving psychological support, thereby improving mental health care and aftercare in oncology [19-21]. As individuals with cancer show elevated levels of distress both during and after cancer treatment, access to (digital) mental health care within this field is of great importance for both prevention and health promotion [4,8].

Despite mixed opinions regarding the software interface, Make It Training was generally rated as user-friendly. The participants most commonly argued for the usability of Make It Training by discussing that high technological literacy was not a requirement for completing the intervention. This finding is consistent with those of previous research showing a link between the use and acceptance of eHealth interventions and users’ technological literacy [23,66,67]. Even though eHealth interventions have the potential to improve health care and aftercare, their implementation often fails because patients face barriers when wanting to make use of these interventions [23,66,67,69]. These barriers include low technological literacy, limitations in technological access, limitations in usability, and limited education in digital advice [69-72]. In addition, there are demographic barriers based on differences in age, socioeconomic status, educational level, language, and culture. Overall, existing barriers to receiving digital interventions due to demographic or structural differences can foster insensitivity within health care [72-74]. Certain individuals with cancer are at risk of being excluded from digital interventions because this population tends to have a higher median age (>60 y) [75], whereas the disease affects individuals with all kinds of demographic characteristics (ie, different cultural backgrounds, socioeconomic statuses, and educational levels). In addition, individuals commonly experience cognitive and physical restrictions during cancer treatment [76]. Thus, for more inclusive health care for individuals with cancer, eHealth interventions need to be designed as barrier free as possible (ie, they should depend less on the user’s technological literacy as well as on other potentially exclusive factors).

Make It Training was compared by the participants to traditional face-to-face therapy even though it was not a specific topic in the interviews. In this regard, Make It Training was described as a helpful intervention, although it was noted that it could not replace traditional psychotherapy. The participants reported the missing therapist interaction as the main reason. In this regard, there was a desire for more therapist interaction within the Make It Training. In addition, a blended therapy format (ie, a combination of the Make It Training with additional face-to-face therapy) was described as the “ideal” format to receive psycho-oncological support. This is in line with previous research supporting the adaptation of blended therapy approaches in psycho-oncology as well [77]. Efficacy research shows that purely self-guided eHealth interventions are associated with smaller effect sizes with a lower completion rate compared to blended therapy interventions, which can be attributed to the missing therapist interaction [78,79]. The results of this study, along with existing research, indicate that it is highly important to adapt eHealth interventions to the patients’ needs [80]. Thus, it is suggested to put emphasis on therapist interaction (ie, blended format) in psycho-oncological eHealth interventions. In this study, a qualitative approach was chosen as we believe that the inclusion of qualitative analyses within efficacy research (ie, the Reduct trial; Bäuerle et al [40]) provides more scientifically sound and transportable results. In this regard, it is important to look beyond surface or aggregate-level evidence to allow for inter- and intrapersonal nuances [81]. These are
often missed in efficacy research but are rather important for a holistic understanding of usefulness in clinical practice [81]. Including qualitative research allows for an investigation of these inter- and intrapersonal nuances as well as for scrutiny of the level of experience, which is an important aspect when evaluating health care interventions such as the Make It Training. Another important strength of this study is the heterogeneity of the sample (ie, all participants were diagnosed with different cancer types and stages), which positively contributed to the generalizability of the evaluation of the Make It Training. In addition, this study provided the research team with information-rich descriptions of the participants’ lived experiences regarding the Make It Training. It was also possible to obtain in-depth feedback on how to design the Make It Training intervention to be more appealing from a patient’s perspective. Practical implications derived from this study are, from patients’ perspectives, the potential of psycho-oncological eHealth interventions such as the Make It Training to improve oncological health care by offering a low-threshold option that provides psychological support independent of time and place and does not interfere with the already time-consuming oncological treatment. However, for routine implementation, they need to be adapted to the patients’ needs and designed to be barrier free and should not require high technological literacy to interact with them. Moreover, even though eHealth interventions do offer efficient psycho-oncological support, they do not replace traditional psychotherapy, and it is suggested to use them as a first-step psychological support in a stepped-care health care approach.

Limitations and Recommendations for Future Research

This study has some limitations. Even though a qualitative approach offers valuable insights into participants’ in-depth experiences, there are limitations regarding qualitative research itself, particularly concerning its generalizability and objectivity [82]. In this study, the decision to use a small sample size might have had a negative impact on the generalizability of the results even though the research team made efforts to select a highly heterogeneous sample. Moreover, a small sample size leads to a smaller data corpus, which can negatively impact the achievement of full thematic saturation. Other limitations include the use of a deductive analysis approach [54] and the risk of selection bias. Moreover, most of the research team members have a background primarily in quantitative methodology. Even though attempts were made to reduce this potential bias by actively involving an expert in qualitative research, this should still be considered a limitation. On the basis of the results of this study, it is suggested that future research put more emphasis on the barrier-free design of interventions and include patients’ perspectives when designing and evaluating eHealth interventions. Moreover, it is suggested that future research investigate blended therapy approaches (ie, a combination of digital psycho-oncological interventions and face-to-face psychotherapy) as this format seems to be appealing for individuals with cancer.

Conclusions

The Make It Training was evaluated as a user-friendly intervention that is helpful for developing functional coping strategies to reduce psychological distress and improve quality of life among individuals with cancer. It has the potential to be implemented as a routine eHealth intervention in oncological health care. Overall, the results of this study, along with the existing literature, support the paradigm shift of including digital mental health care in the treatment of somatic and mental health disorders. e–Mental health interventions such as Make It Training can target both prevention of mental health issues and health promotion and offer a cost-efficient and low-threshold option to receive psycho-oncological support. Moreover, they allow for the retrieval of mental health support content independent of time and place. However, for psycho-oncological eHealth interventions to be actually used, they need to be designed to be barrier free and adapted to the users’ needs.

Acknowledgments

This research was financially sponsored by the Federal Ministry of Education and Research (01KG2021). The authors thank all participants for taking part in this study.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors’ Contributions

JBK contributed to conceptualization, methodology, data curation, writing—original draft preparation, and investigation. FT contributed to conceptualization, methodology, writing—review and editing, and supervision. MT contributed to conceptualization, investigation, writing—review and editing, and supervision. TL contributed to data curation, investigation, and writing—original draft preparation. JH, CS, YE, and JG contributed to conceptualization, investigation, and writing—review and editing. MP contributed to software and writing—review and editing. AB contributed to conceptualization, methodology, writing—original draft preparation, investigation, and supervision.

Conflicts of Interest

None declared.
References


58. Die #1 Software für qualitative Datenanalyse mit der besten KI-Integration. MAXQDA. URL: https://www.maxqda.com [accessed 2023-11-18]


https://cancer.jmir.org/2024/1/e53117

JMIR Cancer 2024 | vol. 10 | e53117 | p.187
(page number not for citation purposes)


Abbreviations

ACT: acceptance and commitment therapy
CBT: cognitive behavioral therapy
CCI: Client Change Interview
COREQ: Consolidated Criteria for Reporting Qualitative Research
Health ITUES: Health IT Usability Evaluation Scale
MBSR: mindfulness-based stress reduction
Differing Content and Language Based on Poster-Patient Relationships on the Chinese Social Media Platform Weibo: Text Classification, Sentiment Analysis, and Topic Modeling of Posts on Breast Cancer

Zhouqing Zhang¹, ME; Kongmeng Liew¹,², PhD; Roeline Kuijer², PhD; Wan Jou She¹,², PhD; Shuntaro Yada¹, PhD; Shoko Wakamiya¹, PhD; Eiji Aramaki¹, PhD

1Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan
2School of Psychology, Speech and Hearing, University of Canterbury, Christchurch, New Zealand
3Department of Information Science, Kyoto Institute of Technology, Kyoto, Japan

Corresponding Author:
Kongmeng Liew, PhD
School of Psychology, Speech and Hearing
University of Canterbury
Private Bag 4800
Christchurch, 8140
New Zealand
Phone: 64 3 369 4333
Email: kongmeng.liew@canterbury.ac.nz

Abstract

Background: Breast cancer affects the lives of not only those diagnosed but also the people around them. Many of those affected share their experiences on social media. However, these narratives may differ according to who the poster is and what their relationship with the patient is; a patient posting about their experiences may post different content from someone whose friends or family has breast cancer. Weibo is 1 of the most popular social media platforms in China, and breast cancer–related posts are frequently found there.

Objective: With the goal of understanding the different experiences of those affected by breast cancer in China, we aimed to explore how content and language used in relevant posts differ according to who the poster is and what their relationship with the patient is and whether there are differences in emotional expression and topic content if the patient is the poster themselves or a friend, family member, relative, or acquaintance.

Methods: We used Weibo as a resource to examine how posts differ according to the different poster-patient relationships. We collected a total of 10,322 relevant Weibo posts. Using a 2-step analysis method, we fine-tuned 2 Chinese Robustly Optimized Bidirectional Encoder Representations from Transformers (BERT) Pretraining Approach models on this data set with annotated poster-patient relationships. These models were lined in sequence, first a binary classifier (no_patient or patient) and then a multiclass classifier (post_user, family_members, friends_relatives, acquaintances, heard_relation), to classify poster-patient relationships. Next, we used the Linguistic Inquiry and Word Count lexicon to conduct sentiment analysis from 5 emotion categories (positive and negative emotions, anger, sadness, and anxiety), followed by topic modeling (BERTopic).

Results: Our binary model ($F_1$-score=0.92) and multiclass model ($F_1$-score=0.83) were largely able to classify poster-patient relationships accurately. Subsequent sentiment analysis showed significant differences in emotion categories across all poster-patient relationships. Notably, negative emotions and anger were higher for the “no_patient” class, but sadness and anxiety were higher for the “family_members” class. Focusing on the top 30 topics, we also noted that topics on fears and anger toward cancer were higher in the “no_patient” class, but topics on cancer treatment were higher in the “family_members” class.

Conclusions: Chinese users post different types of content, depending on the poster–poster-patient relationships. If the patient is family, posts are sadder and more anxious but also contain more content on treatments. However, if no patient is detected, posts show higher levels of anger. We think that these may stem from rants from posters, which may help with emotion regulation and gathering social support.
Introduction

Background

Breast cancer is one of the most common forms of cancer, with an estimated 2 billion people being affected worldwide in 2020 (according to statistics released by the World Health Organization [WHO]), and is consequently a disease familiar to many people. It is a chronic disease with a high mortality rate, which poses a serious threat to human life [1]. For this reason, breast cancer is often viewed negatively, and new diagnoses often trigger sadness, fear, and even psychopathological comorbidities, such as depression [2]. In recent decades, the number of new diagnoses has continued to rise, despite important improvements in medical technologies worldwide [1]. In China, more than 400,000 people were diagnosed with breast cancer in 2020, with approximately 100,000 deaths (according to WHO) [1]. Behind these diagnoses are numerous stories emerging from the experiences of patients or the people around them who are closely intertwined [3]. Therefore, it is not unusual for one to come across discussions on breast cancer in daily life—be it learning about the diagnosis of a loved one or acquaintance or coming across news on a celebrity with breast cancer or even struggling to accept the diagnosis of a close relative. Therefore, a lot of these breast cancer–related narratives take place on social media—lived experiences of people who may have been diagnosed with or who know of someone struggling with breast cancer.

Social media is indispensable in the daily life of billions worldwide; almost everyone is a user of a social media platform [4]. On these platforms, people can share snippets of their lives with other people around them, which double as autobiographical records of their life events. As a social tool, one can smoothly interact and communicate with one’s friends and family over the internet, be it synchronously or asynchronously [5,6]. Such activity leaves digital traces all over the internet, and researchers have since begun using social media posts as resources for uncovering social phenomena [5]. Particularly in the medical field, social media analyses have also been used to great effect, for example, in examining and predicting the epidemiological spread of infectious diseases, such as seasonal influenza and COVID-19 [7,8]. Recently, researchers have also analyzed social media to learn about the perspectives and needs of patients with certain diseases. For example, Kamba et al [9] analyzed a Japanese social media forum (Yahoo Japan) for posts relating to breast cancer and found that the most frequently mentioned concerns pertain to symptoms, screening, and lack of knowledge, to name a few (see also Refs. [10,11]).

However, much of this research has been conducted on Western social media platforms, such as Twitter and Reddit, which have limited penetration in the Chinese market. Chinese internet users have their own social media ecosystems and platforms: Sina Weibo is one of the most widely used and popular social platforms in China and has been called by some as the “Chinese version of Twitter” [12]. Given our research interest in Chinese social media users, we focused our paper specifically on Weibo. As a widely used platform, the number of monthly active users reached 511 million in 2020; Weibo is known by almost everyone in China [13], and posts are known to reflect the diversity of opinions and perspectives by everyday Chinese [14]. Often, users discuss and post about all kinds of topics on Weibo, including topics pertaining to breast cancer. With the large number of users and the diversity of content, Weibo data appear to be a valuable corpus for research on Chinese perspectives from the bottom-up.

Sentiment Analysis on Social Media

To accommodate the large volume of data on the internet, conventional methods, such as qualitative coding, may be too time-consuming and costly. Therefore, modern sociological researchers frequently use computational methods, such as sentiment analysis and topic modeling, to analyze the data. Originating from the field of natural language processing (NLP), sentiment analysis is optimized to deal with the detection and classification of sentiments in (a large number of) texts. By using sentiment analysis, we can infer whether a given text has a positive, negative, or more fine-grained emotional orientation in a given context [15]. In studying social media, researchers analyze the data on social media to obtain public perceptions on a specific topic in contribution to the study and advancement of society [16]. Some researchers have also applied sentiment analysis to measure customers’ needs from their social media posts, thereby obtaining unique insight to improve a brand’s products or services [15]. Researchers have also applied sentiment analysis on social media to predict mental health issues, for example, Wang et al [17] used sentiment analysis to detect users with depression on social networking services.

Regarding breast cancer, sentiment analysis may play a more important role in exploring the patients’ psychological state, such as their perceptions, cognitions, and emotions [18]. Through analyses of tweet sentiments, previous research has confirmed that patients with breast cancer have different polarities (valence) of emotional expression for topics related to breast cancer [19]. For example, support seeking and treatments are associated with positive sentiment, but health care and insurance are associated with negative sentiment. Moreover, posters may not necessarily be patients themselves posting about their experiences or concerns but could be posting about a loved one, a relative, or an acquaintance with breast cancer. Accordingly, posters’ emotional expressions on social media may not only display differences in sentiment, depending on their specified content or aspects (eg, treatment stage or success), but also show differences, depending on their relationship with the patient [20] or if the posters themselves are the patients. In this paper, we define this as the “poster-patient relationship.” Therefore, in studying the usage of social media for emotional expression in the context of breast cancer, sentiment analysis may play an important role in identifying and understanding these differences.
cancer, we propose the necessity to distinguish the poster-patient relationships for each post—whether posts originate from patients themselves or from their friends and relatives or other people.

The Research
Before examining emotional expressions and sentiment, we intended to discern the relationships between poster and patient through the post. Due to the large volume of data, we turned to machine learning for this task. “Machine learning” is the term used to describe both the academic discipline and the collection of techniques that allow computers to undertake complex tasks, and recent advances in machine learning have driven advances in the development of NLP and artificial intelligence (AI) [21]. In NLP, the past 5 years have seen rapid advances in the transformer-based framework, resulting in cutting-edge pretrained language models, such as Bidirectional Encoder Representations from Transformers (BERT) [22], Robustly Optimized BERT Pretraining Approach (RoBERTa) [23], and Generative Pretrained Transformer (GPT)-3 [24], which have greatly improved the effectiveness of downstream tasks (eg, text classification), opening up new avenues for researchers to study society and language [25].

Our aim was to study how users on the Chinese social media platform Weibo post about breast cancer–related topics on social media. Although we took a hypothesis-blind, exploratory approach to data analysis, we focused our discussion on topics surrounding the issue of emotional expression by examining differences in emotional expression, depending on poster-patient relationships. In step 1, we collected data from Weibo and determined poster-patient relationships through 2 stages of classification: first, we identified whether a post references a patient with breast cancer (as opposed to posts that mention breast cancer without naming a specific patient), followed by the poster-patient relationship classification that determined the relationship between the mentioned patient and the author of the post (poster). Ultimately, these 2 stages in step 1 constituted a single classification pipeline to identify poster-patient relationships: whether the post authors are themselves the patients or (1) a family member (family_members); (2) a friend or relative (friends RELATIVES); (3) an acquaintance (acquaintances); (4) from a parasocial relationship, such as a celebrity or public figure (heard relation); or (5) no patient mentioned (no patient). In step 2, we used the LIWC-based dictionary to count the word frequency for each post, with 5 emotional categories (sadness, anger, anxiety, positive, and negative), thereby expanding our target beyond just positive and negative sentiments. Despite the lack of discreet positive emotion categories in the LIWC dictionary, we chose it because it is one of the most widely used and accessible sentiment dictionaries in psycholinguistic research. Next, we used topic modeling to further examine the main topics discussed between each class and how these topics differ across classes. This will allow us to see how social media narratives for patients and posters differ, while shedding light on possible implications for emotional expression via social media.

Methods

Ethical Considerations
As all data used in this study are publicly available and no personal identifiers were obtained, our study was exempt from institutional ethics review. Where applicable, all posts included in this analysis have been paraphrased so that they cannot be traced back to the user. No identifying information (eg, usernames, IDs, or pictures) are included in the main manuscript or in the supplementary material.

Step 1: Poster-Patient Relationship Classification

Data Collection
Since Sina Weibo does not maintain a public application programming interface (API), we used a previously constructed web crawler to request publicly available Weibo posts. Our web crawler simulates a user visiting Weibo’s official website and searches for relevant posts (see the next paragraph for the search procedure). Through this approach, each web search request can obtain up to 50 posts before reinitiating a new search request to retrieve a new set of posts. In our crawler, we were able to set adjustable parameters to specify keywords, the publishing date, location, and interval times between 2 search requests.

We conducted 2 searches with different queries: “breast cancer (乳腺癌)” and “sadness (悲伤)”, as well as “breast cancer (乳腺癌)” and “record (记录)” in Chinese, from January 1, 2018, to December 31, 2021. For both queries, the interval time was set to 15 seconds and the location was unspecified, meaning that we searched for posts from across China. Finally, for the 2 searches with different queries, we obtained 160,182, and 144,125 posts, respectively. For each post, we additionally obtained the user id, username, user type, publish time, post text, location, number of comments, likes, and reposts, which were removed before commencement of analyses.

Next, for the data-cleaning phase, we combined the search results of the 2 queries into a single data set. Duplicate posts were removed through string matching, and obvious advertisements and irrelevant posts were removed by manually checking the data set. This was to ensure the posts were related to narrative accounts pertaining to breast cancer. Finally, this resulted in a cleaned data set containing relevant breast cancer–related narratives from individual users, for a total of 10,322 posts.

Poster-Patient Relationship Classification Criteria
First, we set up 6 categories based on the relationship of the mentioned patient and the author of the post: “post_user,” where the authors are themselves the patients (coded as 0); “family members,” where the authors mention a family member (eg, parent) as the patient (coded as 1); “friends RELATIVES,” where a friend or nonimmediate relative (eg, cousins, aunt) is a patient (coded as 2); “acquaintances,” where a colleague or neighbor (social relationships) is the patient (coded as 3); “heard relation,” where the author may be posting about a celebrity or a famous patient with cancer (coded as 4); and “no patient,” where breast cancer is mentioned generally without being associated with a specific person (coded as 5).
Data Annotation

We randomly portioned 3000 (29.1%) of the 10,322 posts for manual annotation based on the classification criteria, with each data point (post) assigned a label from the 6 aforementioned categories. In the process of labeling, first we determined whether there was a patient in the post (binary classification task), and then we determined whether the poster-patient relationship could be inferred and labeled according to the prespecified classification criteria (multiclass classification task). All data labeling was performed by 1 of the authors who is a native Chinese speaker. See Table 1 for the annotation proportions, and Table S1 in Multimedia Appendix 1 for examples of annotated posts.

To verify that our annotations were objectively labeled and free of subjective bias, we randomly selected 600 (20%) of the 3000 annotated posts, and these were reannotated in the same procedure by another native Chinese annotator who was not part of the research team. Across the 6 categories, the interannotator agreement was good (Cohen χ=0.67) [26], and the original annotations were used to train the classification model.

Table 1. Distribution of annotated posts.

<table>
<thead>
<tr>
<th>Posts</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>No_patient</td>
<td>1089</td>
</tr>
<tr>
<td>Heard_relation</td>
<td>509</td>
</tr>
<tr>
<td>Family_members</td>
<td>443</td>
</tr>
<tr>
<td>Acquaintances</td>
<td>356</td>
</tr>
<tr>
<td>Post_user</td>
<td>338</td>
</tr>
<tr>
<td>Friends RELatives</td>
<td>265</td>
</tr>
</tbody>
</table>

Data Preprocessing

In our study, we chose the pretrained Chinese-RoBERTa-wwm-ext (Chinese RoBERTa) [27] model as our classification model. The Chinese RoBERTa is a large language transformer model based on the RoBERTa architecture [23], trained on a large corpus of the in house–collected extended data containing an encyclopedia, news articles, and web forums, which has 5.4 billion words and is over 10 times bigger than the Chinese Wikipedia [27], and is frequently used for Chinese NLP tasks. To improve the accuracy of the multiclass text classification, we decomposed the classification task over 2 stages (see Ref. [28]): a binary classification task to determine whether a patient was mentioned, followed by a multiclass classifier on posts where a patient was mentioned in order to identify the poster-patient relationship.

The pretrained language model (Chinese RoBERTa) has a limited input character length of 512, and 522 posts in our data set were longer than this character length limit. As such, we used automated text summarization to condense the text length to within 512 characters for these 522 posts using SnowNLP, a Python library that can perform Chinese word segmentation, part-of-speech tagging, sentiment analysis, text categorization, pinyin conversion, traditional simplification, text keyword extraction, text summarization, sentence segmenting, and text similarity estimation [29]. The SnowNLP tool segments posts by sentence and using the TextRank algorithm [30] calculates the weight of each sentence in the post according to the extent to which the content of the sentence represents the content of the text. Finally, all the small units are sorted in reverse order according to their weight scores. When implementing this tool, by setting a number parameter, the corresponding number of sentences is output accordingly, resulting in summarized texts. In Multimedia Appendix 2, we included some examples of automatic summarization.

Classifier Training

Following annotation and data preprocessing, 2 classifiers were constructed for this study in a 2-stage process. In the first stage, a binary classification model was trained to identify whether a patient is mentioned. This was followed by training a multiclass classification model to identify the poster-patient relationship for each post where a patient was mentioned in 1 of 5 classes: post_user, family_members, friends_relations, acquaintances, and heard_relation. This resulted in a total of 6 classes corresponding to the annotations, with the inclusion of the “no_patient” class from the earlier binary classification model. In constructing the 2 classifiers, we specified the task of the RoBERTa model as classification. We monitored the training performance for each epoch through cross-entropy loss. Fine-tuning was implemented under the Pytorch framework, where we used the Amda Optimizer to optimize and update model parameters for training purposes. For testing, sklearn metrics were used to evaluate the binary classification and multiclass classification. In addition, 2400 (80%) of the 3000 annotated posts were used to train the model, and the main parameters for the model training were as follows: batch size=16, learning rate=1.0 × 10⁻⁵, and training epochs=5. We used 600 (20%) posts to test the fine-tuned model.

In the second stage, we removed the “no_patient” class from the annotated data. In total, 1515 (50.5%) posts were used to fine-tune the Chinese RoBERTa model. The main parameters were similar to the binary classifier, with batch size=16, learning rate=1.0 × 10⁻⁵, and training epochs=5. For validation, we used 396 (13.2%) posts to test the trained model.

Step 2: Examining Differences in Emotional Expression

Analysis 1: Sentiment Analysis Based on the LIWC

The LIWC program is a text analysis program that calculates the degree of use for different categories of words across a wide range of text types, from casual conversation to written documents. The LIWC program is based on a comprehensive dictionary of words and phrases that are categorized into different content areas, such as social, personal, and emotional. This allows for the identification of the frequency of words within these categories, providing insight into the emotional expression of the text.

In our study, we used the LIWC program to examine the differences in emotional expression between posts where a patient was mentioned and those where no patient was mentioned. This analysis allowed us to gain a deeper understanding of the emotional content of the posts, which is crucial for understanding the perspectives and experiences of the users.

Analysis 2: Text summarization

Text summarization is a technique used to condense a large amount of text into a shorter, more manageable form while preserving the most important information. This is particularly useful in the context of large datasets, such as the posts from Chinese web forums, where the content can be extensive and difficult to digest. In our study, we used an automatic summarization tool to create summaries of the posts, which helped us to quickly identify the key points and main themes discussed in the posts.

In summary, our study utilized the LIWC program and automatic summarization techniques to analyze the emotional expression and key points of interest in posts from Chinese web forums. This approach provided valuable insights into the perspectives and experiences of the users, and underscored the importance of machine learning and natural language processing tools in understanding and analyzing large volumes of text.
array of texts [31]. This tool was originally developed in English, but researchers have since produced a Chinese version of the LIWC dictionary based on the same criteria [32]. We used an open source Python package to access the Chinese LIWC dictionary. The LIWC dictionary has proved extremely useful in a number of different disciplines and has had a large impact on our understanding of how lexical elements related to cognition, affect, and personal concerns can be used to better understand human behavior [33].

In our study, we focused on the emotion categories to implement the sentiment analysis in our corpus of Weibo posts. We used the LIWC program and its Chinese dictionary to examine 5 emotion categories available in the Chinese LIWC dictionary: positive emotions, negative emotions, sadness, anger, and anxiety. The LIWC dictionary operates by counting the number of terms in each post that corresponds to its internal dictionary for each emotion category, and outputs a score representing the ratio of relevant terms to all identified terms in the post. We then conducted Kruskal-Wallis tests to determine whether positive emotion terms, negative emotion terms, anxiety terms, sadness terms, and anger terms significantly differed between each poster-patient relationship class. If there was a significant effect of the emotion category, we conducted post hoc Dwass-Steel-Critchlow-Fligner (DSCF) pairwise comparisons to compare differences between specific categories.

In this paper, our data are in Chinese, so we had to tokenize our data. We used Jieba for tokenization, which is 1 of the most popular Chinese tokenization tools in NLP [34]. To clean out the noise, we excluded more than 2000 stop words, which were collected from an open source Chinese dictionary of stop words.

### Analysis 2: Topic Modeling

Making sense of a large unstructured corpus through qualitative means is difficult. Therefore, we used topic modeling to better assist us in interpreting data. Topic modeling is a widely used approach to extract common, recurring themes from large amounts of text data through identification and clustering of repeated patterns in words and sentences. In this paper, we adopted the open source BERTopic algorithm [35] to achieve this. BERTopic leverages transformers and class-based term frequency–inverse document frequency (c-TF-IDF) to create dense clusters of words, allowing for easily interpretable topics, while keeping important words in the topic descriptions [35]. Past research [36] has also found that BERTopic-based topic modeling generally yields more theoretically interpretable results than other forms of topic modeling (e.g., latent Dirichlet allocation or Top2Vec). As the BERTopic algorithm only assigns 1 topic to every document (post), we were able to compute topics per class, which allowed uniform comparison of topic distribution for every class (poster-patient relationships), enabling us to observe general trends: which topics are more frequently observed in which class of poster-patient relationship. As long texts are more suitable for modeling and there is no limit to the length of input sentences, during the topic modeling, we replaced the summarized sentences with the original ones. For identified topics, we deliberated on the schema associated with as many words in the topic as possible. Note that this process is largely subjective, so we encourage readers to additionally reference the words contained in each topic, rather than relying solely on the authors’ labels.

In this paper, our data are in Chinese and because the BERTopic model is based on the clustering of individual words to implement topic modeling; therefore, in the process of topic modeling, similar to the sentiment analysis, we needed to tokenize our Chinese data. We again used Jieba for tokenization [34]. To obtain meaningful entities from the topic models, we excluded more than 2000 stop words, which were collected from an open source Chinese dictionary of stop words.

### Results

#### Step 1: Poster-Patient Relationship Classification

**Binary Classifier**

This model was trained to distinguish each post as either mentioning (“patient” class) or not mentioning (“no_patient” class) a patient. We merged the “post_user,” “family_members,” “friends_relations,” “acquaintances,” and “heard_relation” classes into a superordinate “patient” class. The model achieved a high $F_1$-score (see Table 2).

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>$F_1$-score</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>no_patient</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
<td>204</td>
</tr>
<tr>
<td>patient</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>396</td>
</tr>
<tr>
<td>Macro average</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>600</td>
</tr>
</tbody>
</table>

**Multiclass Classifier**

Next, we constructed a multiclass classifier to focus on patient classification: “post_user,” “family_members,” “friends_relations,” “acquaintances,” and “heard_relation.” Results are reported in Table 3.
Table 3. Multiclass classifier's metric report.

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>( F_1 )-score</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>acquaintances</td>
<td>0.76</td>
<td>0.67</td>
<td>0.71</td>
<td>75</td>
</tr>
<tr>
<td>heard_relation</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>102</td>
</tr>
<tr>
<td>family_members</td>
<td>0.93</td>
<td>0.90</td>
<td>0.91</td>
<td>86</td>
</tr>
<tr>
<td>post_user</td>
<td>0.86</td>
<td>0.91</td>
<td>0.89</td>
<td>82</td>
</tr>
<tr>
<td>friends_relatives</td>
<td>0.74</td>
<td>0.84</td>
<td>0.79</td>
<td>51</td>
</tr>
<tr>
<td>Macro average</td>
<td>0.82</td>
<td>0.83</td>
<td>0.83</td>
<td>396</td>
</tr>
</tbody>
</table>

Post Classification

After excluding the annotated data, we were left with 7322 (70.9%) of the 10,322 data points (posts). These posts then underwent the 2-stage classification process. The first stage included a binary classifier to determine whether patient information was identifiable from the post (patient and no_patient), and if a patient was detected, the post then passed to the second stage. This included a multiclass classifier to classify the relationship between the patient and the Weibo poster. In the first stage, 4494 (61.4%) posts were classified as having a patient and 2828 (38.6%) posts as having no patient. Of the former, the relation classifications were as follows (Table 4): the patient was identified as a friend or relative (friends_relatives; \( n=667 \), 14.8%), as the poster (post_user; \( n=705 \), 15.7%), as an acquaintance (acquaintances; \( n=781 \), 17.4%), as a family member (family_members; \( n=961 \), 21.4%), and as someone they had only heard about (heard_relation; \( n=1380 \), 30.7%).

As Tables 1 and 4 show, the rankings of categories by the number of relevant posts were similar regardless of whether the data were manually labeled or predicted by our classifier. The ranking list was no_patient > heard_relation > family_members > acquaintances > post_user > friends_relatives. We noted that the “no_patient” class that did not mention a specific patient was the majority class, which accounted for one-third of the total number of posts (\( n=2828 \), 38.6%). We think that posters use the target words (“breast cancer”) to share some personal thoughts, not necessarily about specific instances of breast cancer or for a targeted patient. Alternatively, they may feel no need to talk about the patient due to the content and style of the post. Except for this class, the distribution of the other poster-patient relationship classes was relatively balanced in the data set.

Table 4. Distribution of predicted posts.

<table>
<thead>
<tr>
<th>Posts, ( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No_patient</td>
</tr>
<tr>
<td>Heard_relation</td>
</tr>
<tr>
<td>Family_members</td>
</tr>
<tr>
<td>Acquaintances</td>
</tr>
<tr>
<td>Post_user</td>
</tr>
<tr>
<td>Friends_relatives</td>
</tr>
</tbody>
</table>

Step 2: Examining Differences in Emotional Expressions

Sentiment Analysis

For subsequent analyses, our aim was to maximize the information we could extract from the data, so manual annotations were combined with the machine-learned predictions for a total of 10,322 posts. We applied the LIWC and the matched Chinese dictionary to count the emotion-related words for each tokenized post. We mainly focused on positive emotion, negative emotion, sadness, anger, and anxiety categories. We calculated the ratio of each emotion category in each post (number of emotion words/number of all tokens). To visualize broad emotional differences among the classified poster-patient relationship classes, we plotted the mean scores for 6 identity categories in each of the 5 emotion categories.

For positive emotions, the “friends_relatives” class had a relatively higher value than the other 5 classes (Table 5). For negative emotions, the “no_patient” class had a relatively higher value than the other 5 classes. For angry terms, the “no_patient” class had a significantly higher value than the other 5 classes, which had almost the same values. For anxiety terms, the “family_members,” “no_patient,” and “post_user” classes had a higher value than the other 3 classes; the “heard_relation” class had the lowest value. For sadness terms, the “family_members,” “no_patient,” and “post_user” classes had a relatively higher value than the other 3 classes.
Table 5. Emotion distribution for each class in the 5 emotion categories (positive emotions, negative emotions, anger, anxiety, and sadness).

<table>
<thead>
<tr>
<th>Emotion Category</th>
<th>no_patient</th>
<th>heard_relation</th>
<th>family_members</th>
<th>acquaintances</th>
<th>post_user</th>
<th>friends_relations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive emotions</strong></td>
<td>0.05567</td>
<td>0.05785</td>
<td>0.05469</td>
<td>0.06581</td>
<td>0.05382</td>
<td>0.07490</td>
</tr>
<tr>
<td><strong>Negative emotions</strong></td>
<td>0.11920</td>
<td>0.09202</td>
<td>0.09933</td>
<td>0.09118</td>
<td>0.09759</td>
<td>0.09386</td>
</tr>
<tr>
<td><strong>Anger</strong></td>
<td>0.01020</td>
<td>0.00490</td>
<td>0.00467</td>
<td>0.00479</td>
<td>0.00469</td>
<td>0.00489</td>
</tr>
<tr>
<td><strong>Anxiety</strong></td>
<td>0.00699</td>
<td>0.00389</td>
<td>0.00674</td>
<td>0.00465</td>
<td>0.00595</td>
<td>0.00430</td>
</tr>
<tr>
<td><strong>Sadness</strong></td>
<td>0.01094</td>
<td>0.00894</td>
<td>0.01107</td>
<td>0.00845</td>
<td>0.01110</td>
<td>0.00928</td>
</tr>
</tbody>
</table>

aNumber of emotion words/number of all tokens.

Next, we statistically examined differences in emotions across poster-patient relationships. Kruskal-Wallis tests showed significant effects for positive emotions (posemo: $\chi^2_5=185.9$, $P<.001$), negative emotions (negemo; $\chi^2_5=156.8$, $P<.001$), anxiety (anx; $\chi^2_5=50.6$, $P<.001$), anger (anger; $\chi^2_5=38.2$, $P<.001$), and sadness (sad; $\chi^2_5=56.8$, $P<.001$). This suggests that for all emotion categories, significant effects were detected across the 6 poster-patient relationship classes. Table 6 reports the post hoc DSCF pairwise comparisons.

Although there were a number of significant effects, here we comment primarily on consistent patterns of results that may be indicative of broader trends in Weibo users with respect to the emotional language used when posting about breast cancer.
We noticed that the “friends relatives” class had significantly higher positive emotions than all other poster-patient relationship classes, and this was followed closely by the “acquaintances” class, which had higher positive emotions than the other remaining poster-patient relationship classes.

Table 6. Pairwise comparisons for the 5 emotion categories.

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 2</th>
<th>Positive emotions</th>
<th>Negative emotions</th>
<th>Anxiety</th>
<th>Anger</th>
<th>Sadness</th>
</tr>
</thead>
<tbody>
<tr>
<td>acquaintances</td>
<td>family_members</td>
<td>W = -7.87</td>
<td>P value &lt;.001 b</td>
<td>W = 3.94</td>
<td>P value .06</td>
<td>W = 7.55</td>
</tr>
<tr>
<td>acquaintances</td>
<td>friends_relative</td>
<td>5.67</td>
<td>&lt;.001 b</td>
<td>1.87</td>
<td>.77</td>
<td>1.29</td>
</tr>
<tr>
<td>acquaintances</td>
<td>heard_relation</td>
<td>-6.75</td>
<td>&lt;.001 b</td>
<td>0.64</td>
<td>0.99</td>
<td>0.91</td>
</tr>
<tr>
<td>acquaintances</td>
<td>no_patient</td>
<td>-10.73</td>
<td>&lt;.001 b</td>
<td>12.13</td>
<td>&lt;.001 b</td>
<td>2.65</td>
</tr>
<tr>
<td>acquaintances</td>
<td>post_user</td>
<td>-8.49</td>
<td>&lt;.001 b</td>
<td>3.13</td>
<td>.23</td>
<td>5.15</td>
</tr>
<tr>
<td>family_members</td>
<td>friends_RELATIVES</td>
<td>13.42</td>
<td>&lt;.001 b</td>
<td>-1.83</td>
<td>.79</td>
<td>-5.90</td>
</tr>
<tr>
<td>family_members</td>
<td>heard_relation</td>
<td>1.78</td>
<td>.81</td>
<td>-3.96</td>
<td>.06</td>
<td>-7.94</td>
</tr>
<tr>
<td>family_members</td>
<td>no_patient</td>
<td>-2.62</td>
<td>.43</td>
<td>8.63</td>
<td>&lt;.001 b</td>
<td>-6.92</td>
</tr>
<tr>
<td>family_members</td>
<td>post_user</td>
<td>-1.08</td>
<td>.97</td>
<td>-0.61</td>
<td>.99</td>
<td>-2.13</td>
</tr>
<tr>
<td>friends_relative</td>
<td>heard_relation</td>
<td>-12.65</td>
<td>&lt;.001 b</td>
<td>-1.54</td>
<td>.89</td>
<td>-0.57</td>
</tr>
<tr>
<td>friends_relative</td>
<td>no_patient</td>
<td>-16.21</td>
<td>&lt;.001 b</td>
<td>9.41</td>
<td>&lt;.001 b</td>
<td>0.94</td>
</tr>
<tr>
<td>friends_relative</td>
<td>post_user</td>
<td>-13.68</td>
<td>&lt;.001 b</td>
<td>1.19</td>
<td>.96</td>
<td>3.69</td>
</tr>
<tr>
<td>heard_relation</td>
<td>no_patient</td>
<td>-5.15</td>
<td>.004 b</td>
<td>14.14</td>
<td>&lt;.001 b</td>
<td>2.01</td>
</tr>
<tr>
<td>heard_relation</td>
<td>post_user</td>
<td>-2.76</td>
<td>.37</td>
<td>2.98</td>
<td>.02</td>
<td>5.02</td>
</tr>
<tr>
<td>no_patient</td>
<td>post_user</td>
<td>1.29</td>
<td>.94</td>
<td>-8.31</td>
<td>&lt;.001 b</td>
<td>3.76</td>
</tr>
</tbody>
</table>

aStandardized Wilcoxon statistic from Dwass-Steel-Critchlow-Fligner (DSCF) pairwise comparisons.
bSignificant P values.

In addition, we found that “no_patient” posts had consistently higher negative emotions than the posts in all other poster-patient relationship classes, but no strong and consistent pattern of difference was observed between other poster-patient relationship classes. This pattern was mirrored strongly in the anger emotion category, suggesting that “no_patient” posts were higher on anger compared to posts in the other poster-patient relationship classes. As “negative emotions” is a broad emotion category containing many other negative emotion words in its dictionary, we think that strong differences observed in anger could be driving the significant difference found in the negative emotions category.

Furthermore, we noticed that with the exception of the “post_user” class, the “family_members” class was generally significantly higher in anxiety than the “acquaintances,” “friends relatives,” “no_patient,” and “heard_relation” poster-patient relationship classes and higher in sadness than the “acquaintances,” “no_patient,” and “heard_relation” poster-patient relationship classes.

Clustered Topics

To gain an overview of why some poster-patient relationship classes were consistently higher in some emotions than other classes, we turned to topic modeling. Using the topics per class function of the BERTopic model, we aimed to compare topical relationships that mirrored some of the identified effects from the sentiment analysis.

We initially found that 139 topics were automatically generated from BERTopic, but this included several topics of low significance, where post counts numbered less than 50. As we wanted to focus on topics of greater relevance, we narrowed our analysis to include only the top 30 (21.6%) topics by topic prevalence across the entire data set, which was sufficient to cover more than 6000 (58.1%) posts. In Table 7 and in Table S2 in Multimedia Appendix 3, we list the top 30 topics with top 30 representative terms and provide a summarized theme for each topic. These are represented by an ID, which represents the ranked prevalence of each topic, while the topic number represents the topic labels assigned for the initial generation. We also visualized the distribution of (poster-patient relationship) classes per topic, which was used to identify topics.
that were more prevalent in a particular class for the analysis. These visualizations are available in our GitHub repository [37].

Table 7. Top 30 terms of top 30 topics from topic modeling.

<table>
<thead>
<tr>
<th>ID</th>
<th>Topic number</th>
<th>Label</th>
<th>Top 30 representative words (Chinese)</th>
<th>Top 30 representative words (translated into English)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Anger</td>
<td>生气,脾气,气死我了,情绪,真的</td>
<td>angry, temper, I’m angry, emotions, really</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Laments</td>
<td>去世,家里,回来,生活,记得</td>
<td>passed away, at home, come back, life, remember</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>Symptoms</td>
<td>乳腺,乳房,肿块,增生,针节,</td>
<td>breast, breast, lump, hyperplasia, node</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>Hospital stays</td>
<td>医生,病人,主任,医院,手术</td>
<td>doctor, patient, director, hospital, surgery</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>Hope and prayers</td>
<td>希望,幸福,生活,人生,幸运</td>
<td>hope, happiness, life, life, lucky</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>Hospitalization</td>
<td>手术,医院,化疗,住院,医生</td>
<td>surgery, hospital, chemotherapy, hospitalization, doctor</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>Lamenting hospitalization</td>
<td>病房,医院,病人,恐惧,患者</td>
<td>ward, hospital, patient, fear, patient</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>Dreams and nightmares</td>
<td>梦里,梦见,到,昨晚,做梦</td>
<td>dream, dreaming, dreaming, last night, dreaming</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>Diagnosis</td>
<td>一年,手术,去看,确诊,希望</td>
<td>a year, surgery, last year, diagnosed, hope</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>Chinese dramas</td>
<td>刘静,女主,男主,欢喜,笑子</td>
<td>Liu Jing, heroine, hero, cheerful, Yingzi</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>School</td>
<td>老师,学生,家长,班主任,上课</td>
<td>teacher, student, parent, classroom, lesson</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>Friends</td>
<td>朋友,闺蜜,离婚,聊天,命理</td>
<td>friend, bestie, divorce, chat, numerology</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>Sleep-wake cycles</td>
<td>熬夜,睡觉,晚上,睡不着,睡着</td>
<td>stay up, sleep, night, sleepless, sleep</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>Passing</td>
<td>去世,消息,难过,死者,刚刚</td>
<td>passed away, news, sad, deceased, just</td>
</tr>
<tr>
<td>14</td>
<td>26</td>
<td>Treatment processes</td>
<td>放疗,化疗,结束,治疗,转移</td>
<td>radiotherapy, chemotherapy, end, treatment, metastasis</td>
</tr>
<tr>
<td>15</td>
<td>33</td>
<td>Treatment effects</td>
<td>治愈,治疗,方案,效果,患者</td>
<td>cure, treatment, protocol, effect, patient</td>
</tr>
<tr>
<td>16</td>
<td>113</td>
<td>Appeal to emotion</td>
<td>开心,心情,事情,难过</td>
<td>happy, mood, things, odds, sad</td>
</tr>
<tr>
<td>17</td>
<td>42</td>
<td>Initiative</td>
<td>面对,压力,生活,健康,人生</td>
<td>face, pressure, life, health, life</td>
</tr>
<tr>
<td>18</td>
<td>11</td>
<td>A Little Red Flower (a popular Chinese movie released in 2020)</td>
<td>小花,一朵,千惠,小红花,病魔</td>
<td>little flower, a, Chie, little red flower</td>
</tr>
<tr>
<td>19</td>
<td>45</td>
<td>Suspicion of breast cancer</td>
<td>怀疑,焦虑症,返祖,胸痛,检查</td>
<td>suspicion, anxiety, revert, chest pain, examination</td>
</tr>
<tr>
<td>20</td>
<td>48</td>
<td>Other cancers</td>
<td>肺癌,肝癌,胃癌,肠癌,吸烟</td>
<td>lung cancer, liver cancer, stomach cancer, bowel cancer</td>
</tr>
<tr>
<td>21</td>
<td>64</td>
<td>Anxiety</td>
<td>焦虑,担心,烦躁,考研,心情</td>
<td>anxiety, worry, irritable, exam, mood</td>
</tr>
<tr>
<td>22</td>
<td>17</td>
<td>Metastasis of cancer cells</td>
<td>转移,癌症,癌细胞,患者,同宏微</td>
<td>transfer, cancer, cancer cells, patient, Yan Hongwei</td>
</tr>
<tr>
<td>23</td>
<td>22</td>
<td>Weibo follows</td>
<td>关注,微博,抗癌,荔枝,记录</td>
<td>concern, microblogging, anti-cancer, lychee</td>
</tr>
<tr>
<td>24</td>
<td>23</td>
<td>Weibo usage</td>
<td>微博,妈妈,努力做到,更新,不想</td>
<td>microblogging, mom, trying to do, update, don't want</td>
</tr>
<tr>
<td>25</td>
<td>85</td>
<td>Treatment side effects</td>
<td>头发,假发,化疗,光头,掉头发</td>
<td>hair, wig, chemotherapy, bald, lose hair</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>Check-up</td>
<td>吉夫,电话,昨天,医生,回去</td>
<td>brother-in-law, phone, yesterday, doctor, go back</td>
</tr>
<tr>
<td>27</td>
<td>63</td>
<td>Female physiology</td>
<td>没事,预防,增生,例假,一去</td>
<td>Nothing, prevention, hyperplasia, period, a go</td>
</tr>
<tr>
<td>28</td>
<td>9</td>
<td>Public figures</td>
<td>陈晓旭,李明,伤官,林黛玉,李婷</td>
<td>Chen Xiaoju, Li Ming, hurt official, Lin Daiyu, Li Ting</td>
</tr>
<tr>
<td>29</td>
<td>58</td>
<td>Treatment stages</td>
<td>化疗,第二次,第三次,结束,白细胞</td>
<td>chemotherapy, second, third, end, white blood cells</td>
</tr>
</tbody>
</table>
Notable Topics

Negative Emotions and Anger
The sentiment analysis suggested that the “no_patient” class had consistently higher negative emotions and anger than all other poster-patient relationship classes. Next, we examined the top 30 topics to identify topics with a similar pattern, which were topics 0, 2, 3, 18, 13, 23, 42, 45, 48, 64, and 113. These spanned a number of overlapping themes. Topic 0, for example, contained terms that directly expressed anger and also appeared to carry the speculation that anger is a cause of breast cancer. Similarly, topics 42, 64, and 113 comprised emotive posts about being positive or hopeful in the face of breast cancer, as well as the anxiety and stress it causes. Posts on topics 3, 48, and 63 contained physiological and medical terms, particularly cancer-related terms, their comorbidities, and their antecedents, and posts on topic 45 appeared to express anxiety at the poster facing a possible cancer diagnosis. Finally, topics 2 and 18 contained posts about the user having a nightmare about breast cancer while sleeping, and topics 13 and 20 were about cancer in everyday life. A guiding theme for these topics is that they seem to relate to the posters’ fears and anger toward cancer in general.

Sadness and Anxiety
Topics 26 and 58 resembled the patterns of relationship classes for sadness and anxiety, in that with the exception of the “post_user” class, the “family_members” class was more prevalent than the other poster-patient relationship classes. These topics shared a common theme, in that they discussed treatment options for breast cancer (eg, chemotherapy, immunotherapy). One explanation could be that immediate family members, as caregivers, were more concerned about breast cancer treatment.

Error Analysis for Machine Learning Classification
Although our classifiers predicted posts well to some extent, we noticed that some cases were mistakenly classified into other categories, according to the metrics from Tables 3 and 6. To explore the possible reasons behind this misclassification, we implemented error analysis.

We found that 1 common reason for these errors was when the patient in a post was unclear and what they said needed to be inferred through semantic understanding. In Table S3 in Multimedia Appendix 4, for example, in post I, the breast cancer patient in the post was the post author (we inferred that the patient should be the poster from reading the post), so according to our classification definition, the true label would be “post_user,” but the predicted label from our classifiers was “acquaintances.” We think that this could be attributed to a mention of a colleague at the beginning of the post and was mistakenly classified into the “acquaintances” class instead. We observed another reason for errors was when the patient was clearly mentioned but there were multiple other actors mentioned in the post as well. Such appearances can greatly affect the classifiers’ prediction. In post II, based on our understanding, the patient appeared to be the poster, but there were many other family members present (eg, father, baby, son, daughter-in-law, granddaughter, grandma). Therefore, post II was mistakenly classified into the “family_members” class instead of the “post_user” class.

Discussion

Principal Findings

Step 1: Poster-Patient Relationship Classification
We fine-tuned the pretrained language model Chinese RoBERTa on our annotations on poster-patient relationships to construct a classification model capable of identifying patients’ relationships with the posters of Weibo posts concerning breast cancer. We subsequently used those classifiers to implement a 2-stage classification process. Both classifiers performed well, and we were generally able to classify poster-patient relationships with moderate-to-high accuracy. This comprised step 1, the poster-patient relationship classification, which was essential to our research question of examining differing Weibo posting styles across poster-patient relationships.

Step 2: Principal Results for Sentiment Analysis and Topic Modeling
In step 2, we used sentiment analysis to compare emotion expressiveness across the 6 poster-patient relationship classes, followed by topic modeling to connect topic content with the emotional difference among identity categories in order to gain an overall understanding. Although this offers only an approximate attempt to interpret the findings of the sentiment analysis, it nevertheless offers an early window into how Weibo posts on breast cancer differ according to the relationship the patient has with the poster. Here, we remind readers that (1) the sentiment analysis was calculated based on broad trends in emotion categories, in that for a specific emotion category, having a higher performance in a relationship class meant that it had a higher frequency across all data, and (2) the distribution of topics per class was performed using the corresponding frequency number of each category across all data, which effectively presented the participation for each relationship class in each topic. In other words, among the 6 relationship classes, the correspondence between each relationship class for each emotion category and the correspondence between each relationship class for each topic can only approximately connect both results to contextualize the emotion from the topic when the relevance is consistent. It does not, however, directly represent the actual relationships between topics and emotion terms, so we caution readers against overinterpreting these results.

Anger and Negative Emotions in “no_patient” Posts
One strong result observed from the sentiment analysis was that “no_patient” posts were consistently higher on anger and negative emotions in general. Considering the topics that are more closely associated with the “no_patient” posts, our interpretation is that posts that omit explicit mentions of patients could indicate the poster’s apprehension, anxiety, or anger toward breast cancer. For example, this could come in the form of a rant. Ranting on social media is a common behavior for expressing stress and dissatisfaction with certain aspects of life. For some users, ranting on a social media platform encourages
social support from other users [38] and is therefore more preferable than ranting in closed media (eg, a diary). Second, ranting on social media has a cathartic effect on the individual with regard to anger reduction [39]. This may thus be a constructive outlet [40] for posters to reduce their negative emotions when feeling particularly angry or anxious toward breast cancer. In these types of posts, we think that the poster may omit explicit mentions of the patient, as these posts are not necessarily of an autobiographical nature but of an expressive nature instead (eg, flow-of-thought writing) and may occur in any situation in which the poster may have a reason to be angry at cancer. For example, posters may be angry at a diagnosis (or prospect) of cancer in themselves or their loved ones, or they may be angry at the problems in society that arise from cancer and associated treatments, which do not necessarily need a target patient.

Sadness and Anxiety in “family_members” Posts
In contrast, sadness and anxiety were consistently higher in posts where close family members (eg, parents) were the patients. This also corresponded with more mentions of treatment options. Past research has documented the significant emotional burden placed on close family members as caregivers of patients with cancer [41]. Moreover, this could be exacerbated by cultural factors: family members are more closely linked to the concept of the self in China, which is largely consistent with interdependent self-construal and collectivistic cultural orientation [42]. In Chinese society, the burden of caregiving often falls to family members, such as adult children [43]. Moreover, (lack of) familial support has been linked to depression and loneliness in elderly Chinese, suggesting the importance of family ties as relational aspects of one’s well-being (eg, interdependent happiness [44]). This may explain the greater mentions of treatment options, and the sadness and anxiety, in Weibo posts where the patient was identified as a family member of the poster; the patient was considered relationally closer and more important to their self-identity, and the poster would also more likely be engaged in caregiving.

This could also be a unique cultural aspect of Chinese individuals. Previous studies have shown that American individuals (elderly) have more independent self-construal, and familial ties, being obligatory, are often less important to the self than friendship ties [44,45]. However, more research is needed to examine similar posts on Western social media platforms for proper cross-cultural examination.

Implications and Future Directions
Our research identified how emotion expression and content change according to the poster’s relationship with the patient, and aligns closely with past research on the stresses and risks family caregivers face for depression and anxiety disorders [36]. This is particularly exacerbated in Chinese culture, where the strain of caregiving is often intensified through cultural norms surrounding filial piety [46]: this means that caregivers often must maintain a patient and positive outlook when interacting with their patients so as not to put an additional burden on the patients. Moreover, discussions about cancer are often seen as taboo in Chinese society, so caregivers cannot easily access social support from their friends and family. However, as social media provides an opportunity for sharing experiences and outreach, it holds immense potential for community building and social support, particularly for familial caregivers (see Ref. [47]). Therefore, we think that social media opens up new opportunities for caregivers (and patients) to seek social support, with reduced fears of breaking social norms and facing judgment from their community. This may even be above and beyond the benefits of social media–based social support in comparatively open Western societies, and we encourage further studies to examine how Chinese internet spaces should be designed to facilitate such social support.

Limitations
To obtain our target data set (long narratives pertaining to breast cancer), we needed to contextualize our initial Weibo queries with additional keywords, in this case “sadness.” Although this enhanced the quality of our data set, it would have biased the data toward more negative sentiments. Nevertheless, despite the overt bias toward negative posts in our sample, significant differences were still observed in poster-patient relationship classes.

During our classification process, we constructed 2 classifiers based on language models. For the binary classifier, the model reached an $F_1$-score of 0.9, and for the multiclass classifier, the model reached an $F_1$-score of 0.8 on average. Although these values are good, there is still some room for improvement for our classifiers. One possibility would be to use a better model for multiclass classification.

In sentiment analysis, we implemented a LIWC-based tool based on the lexical matching of terms for word frequency. Moreover, since only 5 broad affective categories (positive emotions, negative emotions, anger, anxiety, and sadness) were included in this tool, we focused only on these in our study. We think that with newer and more powerful sentiment analysis tools and a larger number of affect categories, the accuracy and granularity of sentiment analysis can be further improved for more valuable insight from the text corpus.

For topic modeling, we used the BERTopic tool to cluster topics, and we found that all the generated topics only had subtle distinctions, which led to several overlaps in similar content among topics. For a better understanding of topics, a qualitative assessment of posts would have yielded deeper insights into the data, but this would not have been practical, given the size of the data set.

Conclusion
In this paper, we studied breast cancer–related narratives on the Chinese social media platform Weibo. Using a pretrained transformer language model (Chinese RoBERTa) as the base model, we fine-tuned 2 models on an annotated subset of the data to classify poster-patient relationships in those posts in a sequential process. Ultimately, we classified all posts according to the identified poster-patient relationships (post_user, family_members, friends_relations, acquaintances, heard_relation, or, if no patient was identified, no_patient).

Next, we implemented sentiment analysis. We used the Chinese LIWC lexicon to examine the sentiment among 6 categories,
focusing on positive emotions, negative emotions, anger, anxiety, and sadness. Through statistical comparisons, we found that emotional expressions present differences among different poster-patient relationship classes. For example, the “no_patient” class had a significantly higher level of anger compared to other classes.

To contextualize these results, we also conducted topic modeling using BERTopic. This showed that posts had different topical content according to the different poster-patient relationships. For example, the “no_patient” class presented more anger in the discussions, while the “family_members” class showed more care for hospitalization and treatment. In sum, our results indicate that patient-poster relationships show differing content and language on Weibo.

Acknowledgments
This work was supported by the Cross-ministerial Strategic Innovation Promotion Program (SIP) on “Integrated Health Care Systems” (grant number JP012425), the JST-Mirai Program (grant number JPMJMI2112), and JST CREST (grant number JPMJCR22N1), Japan. We thank the annotators for their contribution to the data annotation work.

Data Availability
As data contain social media posts that may be linked to individuals, data will be made available upon request to the authors.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Annotation sample.
[DOCX File, 18 KB - cancer_v10i1e51332_app1.docx]

Multimedia Appendix 2
Examples of automatic summarization of long Weibo posts.
[DOCX File, 37 KB - cancer_v10i1e51332_app2.docx]

Multimedia Appendix 3
Top 30 terms of top 30 topics from topic modeling.
[DOCX File, 31 KB - cancer_v10i1e51332_app3.docx]

Multimedia Appendix 4
Examples of error analysis.
[DOCX File, 15 KB - cancer_v10i1e51332_app4.docx]

References


11. Basch CH, Hillyer GC, Yalamanchili B, Morris A. How TikTok is being used to help individuals cope with breast cancer: cross-sectional content analysis. JMIR Cancer 2022 Dec 06;8(4):e42245 [FREE Full text] [doi: 10.2196/42245] [Medline: 36487098]


29. Chen C, Chen J, Shi C. Research on credit evaluation model of online store based on SnowNLP. E3S Web Conf 2018 Sep 14;53:03039. [doi: 10.1051/e3conf/20185303039]


37. zhang9711 / Topic_Modeling_Results. URL: https://github.com/zhang9711/Topic_Modeling_Results [accessed 2024-04-16]


Abbreviations

BERT: Bidirectional Encoder Representations from Transformers
DSCF: Dwass-Steel-Critchlow-Fligner
LIWC: Linguistic Inquiry and Word Count
NLP: natural language processing
RoBERTa: A Robustly Optimized BERT Pretraining Approach
WHO: World Health Organization

©Zhouqing Zhang, Kongmeng Liew, Roeline Kuijer, Wan Jou She, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki. Originally published in JMIR Cancer (https://cancer.jmir.org), 09.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete
bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Evaluating the Quality of Cancer-Related WeChat Public Accounts: Cross-Sectional Study

Peng Pan, PhD; Changhua Yu, MD; Tao Li, BA; Tingting Dai, MD; Hanhan Tian, MD; Yaozu Xiong, MD; Jie Lv, MD; Xiaochu Hu, MD; Weidong Ma, MD; Wenda Yin, MD

Department of Radiation Oncology, The Affiliated Huaian No.1 People’s Hospital of Nanjing Medical University, Huai’an, China

Corresponding Author:
Changhua Yu, MD
Department of Radiation Oncology
The Affiliated Huaian No.1 People’s Hospital of Nanjing Medical University
Radiotherapy Department Building, 4th floor
1 Huanghexilu
Huai’an, 223300
China
Phone: 86 517 8490 7286
Email: 1131603215@qq.com

Abstract

Background: WeChat (Tencent) is one of the most important information sources for Chinese people. Relevantly, various health-related data are constantly transmitted among WeChat users. WeChat public accounts (WPAs) for health are rapidly emerging. Health-related WeChat public accounts have a significant impact on public health. Because of the rise in web-based health-seeking behavior, the general public has grown accustomed to obtaining cancer information from WPAs. Although WPAs make it easy for people to obtain health information, the quality of the information is questionable.

Objective: This study aims to assess the quality and suitability of cancer-related WeChat public accounts (CWPAs).

Methods: The survey was conducted from February 1 to 28, 2023. Based on the WPA monthly list provided by Qingbo Big Data, 28 CWPAs in the WeChat communication index were selected as the survey sample. Quality assessment of the included CWPAs was performed using the HONcode instrument. Furthermore, suitability was measured by using the Suitability Assessment of Materials. A total of 2 researchers conducted the evaluations independently.

Results: Of the 28 CWPAs, 12 (43%) were academic and 16 (57%) were commercial. No statistical difference was found regarding the HONcode scores between the 2 groups (P=0.96). The quality of the academic and commercial CWPAs evaluated using the HONcode instrument demonstrated mean scores of 5.58 (SD 2.02) and 5.63 (SD 2.16), respectively, corresponding to a moderate class. All CWPAs’ compliance with the HONcode principles was unsatisfactory. A statistically significant difference between the 2 groups was observed in the Suitability Assessment of Materials scores (P=0.04). The commercial WPAs reached an overall 55.1% (SD 5.5%) score versus the 50.2% (SD 6.4%) score reached by academic WPAs. The suitability of academic and commercial CWPAs was considered adequate.

Conclusions: This study revealed that CWPAs are not sufficiently credible. WPA owners must endeavor to create reliable health websites using approved tools such as the HONcode criteria. However, it is necessary to educate the public about the evaluation tools of health websites to assess their credibility before using the provided content. In addition, improving readability will allow the public to read and understand the content.

(JMIR Cancer 2024;10:e52156) doi:10.2196/52156

KEYWORDS

cancer; big data; social media; health literacy; WeChat; China; public health

Introduction

According to the International Agency for Research on Cancer (IARC), cancer is the first or second leading cause of premature mortality in over 90 countries. China accounts for 23.7% of global new cases and 30% of deaths. In China, the age-standardized incidence and mortality rates of all cancers were 201.7 per 100,000 and 130.1 per 100,000 in 2018,
respective families [1]. In recent decades, the cancer burden in China has increased, posing a serious threat to public health.

Moreover, with the rapid development of the internet, social media has remarkably changed people’s lifestyles [2]. Similar to Facebook, WeChat, released in 2011 by Tencent Inc, has become the most widely used social networking platform in China, reporting 1299 million users in 2022 [3]. WeChat public accounts (WPA) are application accounts supplied by administrators that can be used for communication and interaction with specific groups via text, pictures, videos, and so forth. Members can follow the WPAs of interest to receive relevant information or messages. In early 2020, more than 1 million papers were posted daily on WeChat [4]. WeChat is one of the most important sources of information for the Chinese public. Pertinently, various health-related information is continuously transmitted among WeChat users. Health-related WeChat public accounts (HWPA) are being rapidly developed [5]. HWPA has an important impact on public health status. Although it is convenient for people to obtain health information from WPAs, the quality of the health information is questionable [6]. Therefore, it is important to evaluate the quality of the HWPA. Previous studies have explored the use of WPAs in health education [7-9]. However, few have focused on the quality of HWPA. Wang et al [10] examined 93 HWPA to evaluate their quality and found that they were substandard according to the Net Foundation Code of Conduct (HONcode) conformity. Furthermore, there is still a lack of general understanding regarding the quality of cancer-related WeChat public accounts (CWPAs). Owing to increasing web-based health-seeking behaviors, the public has become accustomed to obtaining cancer information through WPAs. Therefore, this study was conducted to evaluate the quality of CWPAs. Our study aims to assess the quality and suitability of CWPAs.

Methods

Data Collection

The data used in this study were derived from the Qingbo Big Data platform, the largest third-party evaluation platform for new media in China. Qingbo Big Data Technology Co, Ltd (Beijing, China) was established in October 2014. The company provides big data technology services to the Chinese government, top Chinese news media, and large multinational enterprises [11]. The WeChat communication index (WCI), proposed by Qingbo Big Data, is the most widely used standard for evaluating the influence of WPAs [10]. The WCI comprises 4 primary indicators (the overall spread rate, average spread rate of each paper, title spread rate, and peak spread rate), 8 secondary indicators, and a set of calculation formulas for standardized scores [12]. A higher WCI value indicates a larger WPA influence. We searched for new media in the cancer category of the WPA monthly list (February 1 to 28, 2023) provided by Qingbo Big Data. The CWPAs in the WCI were selected as the survey sample. The exclusion criteria for CWPAs were as follows: (1) having been completed for commercial purposes, and (2) no papers released during the survey period. As a result, 18 CWPAs were excluded according to the criteria. Finally, 28 CWPAs were included in this study (Multimedia Appendix 1). We analyzed 1503 papers released by each CWA on the survey dates.

Evaluation Tools

Quality assessment of the included CWPAs was performed using the HONcode instrument. Health on the Internet is an independent organization that provides health information guidelines for websites based on 8 principles: authoritativeness, complementarity, privacy, attribution, justifiability, transparency, financial disclosure, and advertising policies [13] (Multimedia Appendix 2). The introduction of the HONcode in 1996 was a milestone for web-based health information, as evidenced by the numerous references to the HONcode in the Health Informatics literature. The HONcode has often been used as a major indicator of content accuracy in scientific studies [14]. As CWPAs are used to disseminate cancer-related health knowledge to the public, these WPAs should also comply with the HONcode principles. Thus, we believe analyzing the credibility and reliability of the information on CWPAs using the HONcode instrument is appropriate. We adopted a similar HONcode scoring system to that previously published [15]. For each CWA, the respect or no respect to each HONcode principle was scored as 0 (nonconformity) or 1 (conformity). As a result, the quality of the CWPAs was classified as low (HONcode 0-2), moderate (HONcode 3-5), or high (HONcode 6-8).

The Suitability Assessment of Materials (SAM) created by Doak et al [16] was designed to assess educational material. Applying the SAM can pinpoint specific deficiencies in suitability, and if the material is still in the developmental stage, these deficiencies can be corrected. The SAM comprises 22 criteria in 6 categories: content, literacy demand, graphics, layout, and typography, learning stimulation and motivation, and cultural appropriateness (Multimedia Appendix 3). Within these categories, according to how well they meet the criteria for each item, individual items are rated as follows: not applicable, 0 (not suitable), 1 (adequate), or 2 (superior). The sum of the ratings obtained was divided by the total possible score and transformed into percentages. A total of 3 levels are used to categorize the percentage score: 70%-100%, “superior”; 40%-69%, “adequate”; and 0%-39%, “not suitable” [16]. The SAM has been tested and validated in individuals of various cultural backgrounds [17]. In a study by Chang et al [18], the SAM was proven valid and reliable for evaluating the suitability of health-education materials in Chinese. Therefore, in this study, we used the SAM to evaluate health information released by CWPAs. These CWPAs were classified into academic WPAs and commercial WPAs according to a study by Valizadeh-Haghi et al [19]. The findings of this study revealed that there was a significant association between the website category and the credibility of health websites.

A total of 2 researchers conducted the evaluations. These 2 raters independently evaluated CWPAs’ compliance with the principles of the HONcode and the suitability of the papers released by the CWPAs using the SAM scale. Any controversial assessment results were resolved through real-time negotiations. Cohen κ test assessed interrater reliability, with a score of 0.83 indicating almost perfect agreement [20].

https://cancer.jmir.org/2024/1/e52156JMIR Cancer 2024 | vol. 10 | e52156 | p.206https://cancer.jmir.org/2024/1/e52156 (page number not for citation purposes)
Statistical Analysis
Statistical analysis was performed using SPSS (version 26.0; IMB Corp). Numerical variables are reported as mean (SD) or median (IQR) values. We tested the normality of the distribution of the numerical variables using the Shapiro-Wilk test before proceeding with a parametric or nonparametric test. Parametric variables were compared using the Student t test and nonparametric continuous variables were evaluated with the Mann-Whitney U test. Categorical variables were presented as numbers (n) and percentages (%). Fisher exact test was used to compare categorical variables. Statistical significance was set at P<.05.

Ethical Considerations
According to Article 32 of the ethical review guideline of life science and medical research, which was issued by the National Health Commission of the People’s Republic of China on February 18, 2023, because only publicly available data were involved in our study, the ethical review could be exempted [21]. All data were anonymized.

Table 1. Analysis of cancer-related WeChat public account characteristics by ownership.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Academic CWPAs^a</th>
<th>Commercial CWPAs</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of papers</td>
<td>322</td>
<td>1181</td>
<td>N/A^b</td>
</tr>
<tr>
<td>Number of views, median (IQR)</td>
<td>59,109 (1534.2-60,643.2)</td>
<td>420,237.75 (19,762-439,999.75)</td>
<td>.04</td>
</tr>
<tr>
<td>Number of likes, median (IQR)</td>
<td>226.25 (5.25-231.5)</td>
<td>1640.25 (82.75-1723)</td>
<td>.03</td>
</tr>
<tr>
<td>WCI, mean (SD)</td>
<td>453.57 (255.46)</td>
<td>700.09 (300.20)</td>
<td>.03</td>
</tr>
<tr>
<td>HONcode scores, mean (SD)</td>
<td>5.58 (2.02)</td>
<td>5.63 (2.16)</td>
<td>.96</td>
</tr>
<tr>
<td>SAM^d scores (%), mean (SD)</td>
<td>50.2 (6.4)</td>
<td>55.1 (5.5)</td>
<td>.04</td>
</tr>
</tbody>
</table>

^aCWPA: cancer-related WeChat public account.
^bN/A: not applicable.
^cWCI: WeChat communication index.
^dSAM: Suitability Assessment of Materials.

Health on the Net Foundation Code of Conduct Conformity
The HONcode compliances of the 28 CWPA according to ownership are listed in Table 2. Except for the advertising principle (P=.02), there were no statistical differences regarding the other 7 principles between the academic and commercial groups. CWPA compliance with the HONcode principles was not ideal. Most academic and commercial WPAs failed to meet the principles of transparency and financial disclosure (58% vs 62% and 58% vs 69%, respectively). One-third of academic (4/12, 33%) and almost half of the commercial WPAs (7/16, 44%) did not respect this attribution principle. All academic WPAs and 94% (n=15) of the commercial WPAs received a full score on the justifiability principle. Compliance was also uneven for authoritative and complementarity principles. A greater proportion of academic WPAs achieved full scores in the authoritative and complementarity principles (10/12, 83% and 11/12, 92%, respectively), compared with commercial WPAs (11/16, 69%). Finally, only 25% (n=3) of academic WPAs received a full score in the advertising policy principle, compared with commercial WPAs (12/16, 75%); that is, three-quarters of the academic WPAs did not clearly distinguish advertising from editorial content.
Table 2. Evaluating HONcode scores according to ownership.

<table>
<thead>
<tr>
<th>HONcode principles</th>
<th>Proportion of academic CWPA(^a) with full score (n=12), n (%)</th>
<th>Proportion of commercial CWPA with full score (n=16), n (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authoritative</td>
<td>10 (83)</td>
<td>11 (69)</td>
<td>.66</td>
</tr>
<tr>
<td>Complementarity</td>
<td>11 (92)</td>
<td>11 (69)</td>
<td>.19</td>
</tr>
<tr>
<td>Privacy</td>
<td>12 (100)</td>
<td>16 (100)</td>
<td>N/A(^b)</td>
</tr>
<tr>
<td>Attribution</td>
<td>8 (67)</td>
<td>9 (56)</td>
<td>.71</td>
</tr>
<tr>
<td>Justifiability</td>
<td>12 (100)</td>
<td>15 (94)</td>
<td>.38</td>
</tr>
<tr>
<td>Transparency</td>
<td>5 (42)</td>
<td>6 (38)</td>
<td>.82</td>
</tr>
<tr>
<td>Financial disclosure</td>
<td>5 (42)</td>
<td>5 (31)</td>
<td>.69</td>
</tr>
<tr>
<td>Advertising policy</td>
<td>3 (25)</td>
<td>12 (75)</td>
<td>.02</td>
</tr>
</tbody>
</table>

\(^a\)CWPA: cancer-related WeChat public account.
\(^b\)N/A: not applicable.

Suitability of Papers From WPAs

Table 3 presents the analysis of the CWPA’s readability using the SAM. Among the 6 categories, a statistically significant difference was found between academic and commercial WPAs in literacy demand (P=0.02). In most cases, the mean scores of academic WPAs were lower than those of commercial WPAs, except for the cultural appropriateness items. However, no statistical differences were found between the 2 groups regarding content (P=0.53), graphics (P=0.07), layout and typography (P=0.84), learning stimulation and motivation (P=0.95), or cultural appropriateness (P=0.78). None of the CWPA achieved a superior score on the SAM items. The percentages of criteria met in each of the 6 SAM categories ranged from the lowest for learning stimulation and motivation to the highest for content.

Table 3. Evaluating Suitability Assessment of Materials scores of papers on the cancer-related WeChat public account according to ownership.

<table>
<thead>
<tr>
<th>SAM(^a) items</th>
<th>Academic CWPA(^b), mean (SD)</th>
<th>Commercial CWPA, mean (SD)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content (purpose is evident, content regarding behavior, scope is limited, and summary or review included)</td>
<td>5.08 (0.51)</td>
<td>5.25 (0.86)</td>
<td>.53</td>
</tr>
<tr>
<td>Literacy demand (reading grade level, writing style, active voice, vocabulary uses common words, context is given first, and learning aids via “road signs”)</td>
<td>4.58 (0.79)</td>
<td>5.56 (1.15)</td>
<td>.02</td>
</tr>
<tr>
<td>Graphics (cover graphic shows purpose; type of graphics; relevance of illustrations; list, tables, etc explained; and captions used for graphics)</td>
<td>4.67 (1.15)</td>
<td>5.63 (1.41)</td>
<td>.07</td>
</tr>
<tr>
<td>Layout and typography (layout factors, typography, and subheads used)</td>
<td>3.83 (0.58)</td>
<td>3.88 (0.50)</td>
<td>.84</td>
</tr>
<tr>
<td>Learning stimulation and motivation (interaction used, behaviors are modeled and specific, and motivation [self-efficacy])</td>
<td>1.67 (0.98)</td>
<td>1.69 (0.79)</td>
<td>.95</td>
</tr>
<tr>
<td>Cultural appropriateness (match in logic, language, and experience; cultural image; and examples)</td>
<td>2.25 (0.62)</td>
<td>2.19 (0.54)</td>
<td>.78</td>
</tr>
</tbody>
</table>

\(^a\)SAM: Suitability Assessment of Materials.
\(^b\)CWPA: cancer-related WeChat public account.

In the content category, most WPAs stated their purpose in the titles and contained related information within the necessary scope. However, some (8/28, 28.6%) of WPAs contained behavior-related context in presenting content; that is, the content mainly included facts about cancers and not guides for readers’ behavior or decision-making. Additionally, 14% (n=4) of WPAs did not include a summary or review. Regarding literacy demand, 1 WPA was classified as not suitable, 26 as adequate, and only 1 was superior. The cover graphics for most papers released by the WPAs were rated as superior. However, 68% (n=19) of the WPAs used illustrations inappropriately. Moreover, one-quarter (7/28, 25%) of the papers did not include captions that detailed the information in the tables and graphs. The layouts of most papers received high scores; for example, most were adequate or superior regarding typography and font size. Lower ratings were caused by the tendency to include too much information under the subheadings. The learning stimulation and motivation categories had the lowest ratings. None of the studies provided web-based learning stimulation. The content for behavioral modeling and self-efficacy of 93% (n=26) of the WPAs was adequate or not suitable. Most WPAs (26/28, 93%) were rated as adequate or superior for using...
positive images and examples for the cultural appropriateness category.

Discussion

Principal Findings

WeChat is the most popular platform for acquiring health information. Health information acquisition via WeChat is more convenient, timely, and cost-effective; moreover, it protects privacy and avoids embarrassment. Furthermore, the technical development of big data and the Internet of Things allows individuals to access, track, and customize health information. To a certain extent, WeChat contributes to greater freedom regarding individual health decisions.

The public encounters problems through the internet via an overload of information. In our study, searching for “cancer” in the Qingbo search engine generated 46 WPAs. Worryingly, the information presented in these WPAs is not sufficiently credible; that is, valid and valuable information is obscured by irrelevant and misleading information. To our knowledge, this was the first study to evaluate the quality and readability of WPAs concerning cancer. Our study mirrored other studies’ findings on various topics [22-24].

The HONcode instrument for health-related web resources has been available for 20 years. A failure to comply with the HONcode criteria indicates that users may encounter websites that are not sufficiently reliable. These websites may contain inaccurate, misleading, and inadequate information, which can influence preventive actions and decision-making regarding cancer treatment choices.

This study’s findings revealed that all CWPAs’ compliance with the HONcode principles was unsatisfactory. Although one could intuit that the information found in academic CWPAs would yield the highest quality information, our study found this was not always true. Specifically, there was no statistical difference between academic WPAs and commercial CWPAs regarding HONcode sum scores and most HONcode categories. Thus, academic institutions must take substantial steps to improve the credibility of their WPAs to comply with the HON principles.

Compliance with the authority criterion reflects the credibility of the information source because this principle proves that the information provided by experts is reliable [25]. In this study, one-quarter (7/28, 25%) of the surveyed CWPAs did not specify the names or expertise of the authors. In a similar study evaluating Persian language health websites on Ebola, the authorities obtained the lowest score [22]. While the public needs sufficient information about the author’s identity to assess the trustworthiness of information, CWPAs must pay more attention to this criterion to increase trustworthiness for their readers. The complementarity aspect of web-based medical information should be clearly stated on health websites as such information is intended to provide support and training for readers and should not be a substitute for direct medical advice [19]. Nevertheless, 21.4% (n=6) of the surveyed CWPAs did not consider this criterion, which may have led to misuse of information. Moreover, CWPAs should describe their privacy policies and define how they handle users’ private information such as email addresses and content. This policy is among the 7 core issues in website usability design and is particularly important for creating effective websites [26]. Satisfactorily, all CWPAs assessed in this study identified their privacy policies. According to the attribution principle, the publication date and most recent content updates should be posted on the website. Adherence to this principle can ensure the credibility of health websites. This study revealed that the attribution principle was considered in more than half (17/28, 61%) of the CWPAs. Nevertheless, 39% (n=11) of the CWPAs did not pay sufficient attention to this principle. The justifiability criterion indicates that any information on a website must support claims regarding the benefits or performance of a particular treatment, medication, or medical device. Overall, in this study, the adherence to the justifiable principle was good. The transparency principle states that when additional information is required, people must be able to connect with content editors and communicate with webmasters. Unfortunately, based on the present findings, this principle was only considered in 39% (n=11) of the CWPAs. Financial disclosure and advertising principles imply that there should be a clear distinction between commercial and scientifically edited content presented on CWPAs. If advertising is a source of funding for a WPA, the financial disclosure policy for presenting such content should be clearly stated. Moreover, failure to comply with advertising policies indicates that individuals may be unable to distinguish advertisement information from the main content. Access to such WPAs may guide readers toward unreliable information that may threaten their health. However, only 36% (n=10) of the surveyed CWPAs considered financial disclosure principles. More importantly, 75% (n=9) of the academic WPAs failed to comply with the advertising policy. In contrast, only 25% (n=4) of commercial WPAs failed to comply with this principle.

The content provided to the public must not only meet the reliability standard but also be at the required reading level that allows people of all educational levels to understand and process information related to their disease and treatment options [27]. Notably, lower overall health literacy is associated with increased complications, hospitalizations, poor understanding of the disease, and increased health care costs [28]. This study determined the CWPAs’ information suitability to be “adequate.” Cultural and linguistic differences inevitably lead to differences in people’s health-related behaviors and understanding of web-based health information. Thus, website owners must consider additional acculturation factors when publishing health information. However, this creates higher requirements for user cultural literacy [11].

Regarding scoring dimensions, most papers published by the CWPAs had appropriate cover pictures and attractive titles that clearly described the paper’s purpose, a good layout and typography, and were culturally suitable. However, the nonstandard use of charts and the lack of charts used as illustrations were common problems. More than half of the papers included pictures with weak relevance to the content of the papers or even harmful overstatements and stereotypical cultural characteristics. In addition, regarding vocabulary,
readers had difficulty reading papers generated by professional WPAs because they usually use more scientific terms.

The cultural appropriateness of health-education materials is enhanced when readers view illustrations and graphics that are easily recognizable and depict people similar to themselves and those around them. Many factors affect health care including cultural beliefs and practices [17]. Thus, it is important to consider these factors when designing health education materials.

We are particularly interested in examining the facilitation of self-efficacy. Applying the concept of self-efficacy is an effective means of promoting positive health behaviors and informed decision-making. Self-efficacy theory explains and predicts how people influence their motivation and behavior; to enhance self-efficacy, materials must model the desired behavior using someone similar to the intended audience [17]. Very few of the reviewed materials used appropriate methods to enhance readers’ self-efficacy.

Based on our findings, although the suitability of health information released by CWPAs was at a moderate level, the overall quality of accessible information on CWPAs was inadequate. Failure to comply with all HONcode criteria in these CWPAs shows that while searching for WPAs, users will encounter impressive websites, and consequently, low-quality information that can affect their health care practices for cancer. Reliable and readable information is essential for overcoming the potential negative aspects of web-based health information. Providing information in shorter sentences with simple words and using figures or videos may help improve the public’s understanding of cancer and cater to people with varying levels of health literacy. This highlights the importance of understanding the quality of CWPAs by providers and guiding the public toward reliable sources. Finally, it is recommended that the papers of CWPAs be subjected to some form of peer review, similar to those used for journal paper submissions, before the final upload. This would create a core set of high-quality, publicly available information.

Limitations

This study has some limitations. The study was conducted between February 1 and 28, 2023; therefore, it does not completely and comprehensively represent other studies conducted at different times. However, owing to the dynamic characteristics of the web, search results vary at different times and places. New websites are constantly being created, while some websites are being disbanded. Second, there are many evaluation indices for WPAs; however, horizontal comparisons of these indices are lacking. We chose the WCI proposed by Qingbo Big Data as the ranking basis for the influence of WPAs, which may have resulted in selection bias. Finally, this study was conducted only on Chinese websites. Therefore, the results of this study may differ from those conducted in other languages.

Conclusions

This study revealed that CWPAs are not sufficiently credible. WPA owners must endeavor to create reliable health websites using approved tools such as the HONcode criteria. However, it is necessary to educate the public about the evaluation tools of health websites to assess their credibility before using the provided content. In addition, improving readability will allow the public to read and understand the content.

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors’ Contributions

PP and CY participated in the conceptualization of the paper. TD, HT, XH, and WM conducted the data searches on the internet. JL and WY conducted data evaluation. YX performed statistical analysis. TL critically reviewed the manuscript for important intellectual content. PP structured and wrote the paper. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Raw data of the cancer-related WeChat public accounts.
[XLSX File (Microsoft Excel File), 16 KB - cancer_v10i1e52156_app1.xlsx ]

Multimedia Appendix 2
The HONcode principles.
[DOCX File, 17 KB - cancer_v10i1e52156_app2.docx ]

Multimedia Appendix 3
The Suitability Assessment of Materials criteria and descriptions.
[DOCX File, 20 KB - cancer_v10i1e52156_app3.docx ]
References

3. 2022 Second quarter results presentation. Tencent Holdings Limited. URL: https://static.wtcn.com/uploads/2022/08/17/d0d74555defe43fe57a9a3e05a06e87b.pdf [accessed 2022-08-17]


Abbreviations

- CWPA: cancer-related WeChat public account
- HWPA: health-related WeChat public account
- IARC: International Agency for Research on Cancer
- SAM: Suitability Assessment of Materials
- WCI: WeChat communication index
- WPA: WeChat public account

©Peng Pan, Changhua Yu, Tao Li, Tingting Dai, Hanhan Tian, Yaozu Xiong, Jie Lv, Xiaochu Hu, Weidong Ma, Wenda Yin. Originally published in JMIR Cancer (https://cancer.jmir.org), 30.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Original Paper

Engagement With a Relaxation and Mindfulness Mobile App Among People With Cancer: Exploratory Analysis of Use Data and Self-Reports From a Randomized Controlled Trial

Sonja Schläpfer¹, MSc; Fabian Schneider², BSc; Prabhakaran Santhanam², MSc; Manuela Eicher³,⁴, Prof Dr; Tobias Kowatsch²,⁵,⁶, Prof Dr; Claudia M Witt¹, MBA, Prof Dr Med; Jürgen Barth¹, PhD

¹Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland
²Centre for Digital Health Interventions, Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland
³Institute of Higher Education and Research in Healthcare, Faculty of Biology and Medicine, University of Lausanne and Lausanne University Hospital, Lausanne, Switzerland
⁴Department of Oncology, Lausanne University Hospital, Lausanne, Switzerland
⁵Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland
⁶School of Medicine, University of St.Gallen, St.Gallen, Switzerland

Corresponding Author:
Sonja Schläpfer, MSc
Institute for Complementary and Integrative Medicine
University Hospital Zurich and University of Zurich
Sonneggstrasse 6
Zurich, 8091
Switzerland
Phone: 41 44 255 94 51
Email: sonja.schlaepfer@usz.ch

Abstract

Background: Mobile health (mHealth) apps offer unique opportunities to support self-care and behavior change, but poor user engagement limits their effectiveness. This is particularly true for fully automated mHealth apps without any human support. Human support in mHealth apps is associated with better engagement but at the cost of reduced scalability.

Objective: This work aimed to (1) describe the theory-informed development of a fully automated relaxation and mindfulness app to reduce distress in people with cancer (CanRelax app 2.0), (2) describe engagement with the app on multiple levels within a fully automated randomized controlled trial over 10 weeks, and (3) examine whether engagement was related to user characteristics.

Methods: The CanRelax app 2.0 was developed in iterative processes involving input from people with cancer and relevant experts. The app includes evidence-based relaxation exercises, personalized weekly coaching sessions with a rule-based conversational agent, 39 self-enactable behavior change techniques, a self-monitoring dashboard with gamification elements, highly tailored reminder notifications, an educational video clip, and personalized in-app letters. For the larger study, German-speaking adults diagnosed with cancer within the last 5 years were recruited via the web in Switzerland, Austria, and Germany. Engagement was analyzed in a sample of 100 study participants with multiple measures on a micro level (completed coaching sessions, relaxation exercises practiced with the app, and feedback on the app) and a macro level (relaxation exercises practiced without the app and self-efficacy toward self-set weekly relaxation goals).

Results: In week 10, a total of 62% (62/100) of the participants were actively using the CanRelax app 2.0. No associations were identified between engagement and level of distress at baseline, sex assigned at birth, educational attainment, or age. At the micro level, 71.88% (3520/4897) of all relaxation exercises and 714 coaching sessions were completed in the app, and all participants who provided feedback (52/100, 52%) expressed positive app experiences. At the macro level, 28.12% (1377/4897) of relaxation exercises were completed without the app, and participants' self-efficacy remained stable at a high level. At the same time, participants raised their weekly relaxation goals, which indicates a potential relative increase in self-efficacy.

Conclusions: The CanRelax app 2.0 achieved promising engagement even though it provided no human support. Fully automated social components might have compensated for the lack of human involvement and should be investigated further. More than...
one-quarter (1377/4897, 28.12%) of all relaxation exercises were practiced without the app, highlighting the importance of assessing engagement on multiple levels.

*(JMIR Cancer 2024;10:e52386)* doi:10.2196/52386

**KEYWORDS**

mobile health; mHealth; digital health; eHealth; smartphone; mobile phone; implementation; adherence; self-guided; unguided; fully automated; conversational agent; chatbot; behavior change; tailoring; self-care; cancer; app development

**Introduction**

**Background**

Mobile health (mHealth) apps offer unique opportunities to deliver self-care interventions and support behavior change, but poor user engagement and retention rates pose substantial challenges. mHealth apps are a convenient approach to facilitate behavior change with the potential to reach large numbers of people [1-3]. However, in the same manner that mHealth apps provide easy access with a low barrier to start an intervention, they also provide a low barrier to stop using an intervention, turning a great advantage of mHealth apps into a fundamental challenge [2]. Low engagement is problematic because mHealth apps that support healthy behaviors can only be effective if people take an active role, learn the necessary skills to change their behavior, and apply the skills to everyday life, making engagement a pivotal prerequisite to health behavior change [4-8]. In studies using mHealth apps, poor engagement can also confound the outcome and impact the validity of the results as study dropouts may differ from completers [2,9]. While many mHealth apps have significant issues with sustained engagement [10-17], this is particularly true for fully automated mHealth apps without any human support, also termed unguided or self-guided mHealth apps. A high level of human support in guided mHealth apps is typically associated with better engagement rates but at the cost of reduced scalability [18,19]. Hence, to increase the effectiveness of behavior change interventions and improve mHealth studies, it is critical to better understand what makes people stay engaged with mHealth apps [2,20-22] and especially with fully automated mHealth apps as the latter are more likely to be disseminated widely [23].

User engagement has been conceptualized differently across disciplines, but there is a consensus that engagement with an mHealth app needs to be examined on different levels [7,24]. The different levels stem from the crucial distinction between moment-to-moment engagement with the intervention at the micro level (ie, app use and user experience) and engagement with the broader intervention goal at the macro level (ie, target behavior) [5,7]. The micro and macro levels are closely interlinked, and engagement at the different levels can vary over time [5]. For example, during the initial use phase of mHealth apps, moment-to-moment engagement with the app may serve as preparation for behavior change. In a later phase, when people apply the skills they learned to everyday life, use of the app may no longer be required for engagement with the targeted behavior. Hence, reduced app use could be a sign of success rather than failure [2], highlighting the importance of comprehensively assessing engagement.

Most mHealth studies assess engagement with system use data at the micro level but do not consider engagement measures at the macro level. At the micro level, system use data such as the number of log-ins or the amount and type of content used are frequently applied as the only measure of engagement with an mHealth app. However, although system use data undoubtedly provide valuable information on certain aspects of microlevel engagement, these data are not considered a valid measure of engagement on their own [24]. Greater efforts are needed to combine different data sources, such as pairing system use data with self-report data or qualitative methods, to better understand the user experience [5,6,17,24]. At the macro level, assessing engagement remains a challenge and is often neglected in mHealth studies. To support research in this area, recent reviews have provided a valuable overview of available measures for exploring engagement in the behavior change process in daily life [5,6,24]. The listed measures to assess macrolevel engagement include sensor data to track behavior in real-life settings, analysis of social media patterns, and the repeated assessment of psychological constructs that are hypothesized to be important determinants of behavior change (eg, self-efficacy) [24]. Changes over time in psychological constructs such as self-efficacy could indicate engagement in the behavior change process [24]. Given the complexity of engagement as a construct, other measures of macrolevel engagement might be useful depending on the specific research context. Thus far, little research has been conducted applying these or other measures at the macro level of engagement and exploring their use in an mHealth behavior change setting [24].

**Objectives**

We examined engagement at both a micro and a macro level with a newly developed relaxation and mindfulness app to reduce distress in people with cancer (CanRelax app 2.0) within a fully automated randomized controlled trial (RCT) over 10 weeks. The CanRelax app 2.0 is based on a first app version piloted in a feasibility study [25] and now includes more relaxation resources, a conversational agent, gamification elements, and 39 behavior change techniques (BCTs) translated into designed app features. The aims of this paper were to (1) describe the theory-informed development of the CanRelax app 2.0, (2) describe engagement with the app over 10 weeks as total app use and user feedback (micro level) and as self-efficacy and reported relaxation practices without using the app (macro level), and (3) examine whether engagement was related to user characteristics.
Methods

Study Design

The presented data originated from a larger RCT with an additional nonrandomized third arm. The study aimed to evaluate the effectiveness of the CanRelax app 2.0 in reducing distress in people with cancer who experience high distress compared with a waitlist control group. The primary end point was distress after 10 weeks assessed using the Patient Health Questionnaire Anxiety and Depression Scale [25]. Secondary outcomes were well-being (5-item World Health Organization Well-Being Index [26]), self-regulation (Multidimensional Assessment of Interoceptive Awareness Self-Regulation subscale [27]), and the course of distress over time (4-item Patient Health Questionnaire [28] and Distress Thermometer [29]; Multimedia Appendix 1 [25-31]). Eligible participants who self-reported high distress at baseline (Distress Thermometer score of ≥5 [29]) were randomized using 1:1 block randomization stratified by sex; those who self-reported low distress at baseline (Distress Thermometer score of <5 [29]) were included in a third arm as a nonrandomized intervention group to further explore user engagement. This nonrandomized intervention group received immediate access to the app (the same app as the randomized intervention group); the waitlist control group received full access to the app after 10 weeks. All groups were allowed to continue usual care and other interventions (including self-care interventions) as needed. As per sample size calculation, the target sample size was 210 participants in the randomized study arms (105 per arm); the sample size was not predefined for the nonrandomized third arm. The study was registered a priori at the German Clinical Trials Register (DRKS00027546; registration date: February 23, 2022). For this paper, data were taken from participants randomly assigned to the intervention group and participants assigned to the nonrandomized third arm. Further information on the study design and assessments is provided in Multimedia Appendix 1 [25-31]. The results of the RCT will be reported elsewhere.

Inclusion Criteria

People were eligible to participate in the study if they (1) had received a cancer diagnosis within the last 5 years regardless of the type of cancer or stage at diagnosis, (2) were aged ≥18 years, (3) were fluent in German, (4) had a smartphone with regular internet access, and (5) gave informed consent to participate in the study. The exclusion criteria were suicidal ideation and known pregnancy according to participants’ self-reports. For this study, we analyzed an exploratory sample of the first 100 study participants who received full access to the CanRelax app 2.0 at inclusion. This corresponds to the sample needed to detect a meaningful difference (effect size d=0.8) in engagement between subgroups (high, 67/100, 67%; vs low, 33/100, 33%; distress) with a power of 0.95 (α=.05). The study was advertised for distressed individuals with cancer. Hence, we expected more high-distress than low-distress participants and assumed a ratio of approximately 2:1. Participants were excluded from the analysis if they withdrew from the study and requested that we exclude their data. In these cases, we included the next participant who received full access to the app at inclusion so that we had data from 100 participants for analysis.

Recruitment Procedure

We launched the app in July 2022 through the Apple App Store and Google Play Store in Switzerland, Germany, and Austria. At the same time, we established a project website to facilitate recruitment. The website presented a summary of the study with key information such as the eligibility criteria, pictures of the app, and audio samples. It also included QR codes containing web links to the CanRelax app 2.0 in both app stores. We used social media sites (ie, Facebook, Twitter, and LinkedIn) and more traditional approaches (eg, consultations with health care providers, printed flyers, newsletters, and a press release by the University Hospital Zurich) to recruit study participants. Interested individuals could download the app free of charge and start by completing the app onboarding process as a first introduction to the app and the study. From the beginning, users were explicitly informed that they were interacting with a conversational agent, not a person. All study processes were fully automated; screening questions, study information and consent, enrollment, data collection, and all steps up to completion of follow-up were managed entirely through the CanRelax app 2.0. Participants had no contact with the research team at any time during the study unless they contacted the research team to ask questions before consenting or in case of technical issues. The RCT completed recruitment successfully in February 2023. Data collection was ongoing at the time of writing this paper.

Intervention

Overview

The intervention was a fully automated mHealth app designed specifically to improve distress in adults with cancer through one type of self-care behavior (relaxation). Participants had access to the CanRelax app 2.0 over 20 weeks (10 weeks of intervention and 10 weeks of follow-up). On day 1, participants selected an outcome goal from a 5-item list in the app, including “find inner peace” (default if no choice was made), “improve coping strategies;” “build self-confidence,” “increase joy in life,” and “just curious.” Participants were periodically reminded of this goal during the intervention, and it was displayed in the dashboard of the app. During the intervention, participants could also set weekly relaxation goals in terms of a targeted number of relaxation exercises per week (with 1 exercise per week at minimum and a default of 3 exercises per week irrespective of the type of exercise). Weekly coaching sessions with a text-based conversational agent called Lumy provided motivational input for effective and lasting behavior change (integration of relaxation into daily life). Participants were encouraged to set small, realistic relaxation goals for themselves, choose and practice any relaxation exercise at their convenience to meet their goals, and chat with Lumy each week. The minimum expectation for participation in this intervention was completing at least one relaxation exercise and one coaching session per week over the 10-week intervention period.
Technical Implementation of the CanRelax App 2.0

The app was built using MobileCoach (version 21.9.1), an open-source software platform for digital biomarker and health intervention research [32,33]. Conceptually, the app implements the Talk-and-Tools paradigm, which was applied successfully in the domain of mHealth behavior change interventions [34]. The app offers a user interface with a conversational agent (the talk) and a broad range of tools (Multimedia Appendix 2). Our conversational agent Lumy is visually represented by a neutral (nonhuman) avatar (Multimedia Appendix 3). By choosing a nonhuman avatar, we aimed to create an inclusive experience for all app users and followed best practices and design principles of popular commercial mindfulness and relaxation apps (eg, Headspace). The tools include evidence-based relaxation exercises, a self-monitoring dashboard with metrics on participants’ goals and progress, an educational video clip, personalized in-app letters, frequently asked question (FAQ) sections on the mechanisms and benefits of relaxation as well as on creating healthy habits, and tailored reminder notifications to support regular relaxation practice and engagement with the app. Screenshots of the app can be found in Figures 1-3.

Figure 1. Screenshot of the CanRelax app 2.0—resource library with relaxation exercises. (1) Filter for exercise characteristics (male or female voice with or without background music), (2) search results (can be scrolled for further exercises), (3) audio files, and (4) breathing training.
Figure 2. Screenshot of the CanRelax app 2.0—interaction with the conversational agent Lumy (reviewing and adjusting goals). (1) Lumy: “Well done, Robin. Now let’s talk about the goal you want to set for yourself in the coming weeks.” (2) Answer options: “Okay” or “I prefer to skip this part.”

Figure 3. Screenshot of the CanRelax app 2.0—dashboard. (1) Intervention start date, current week, and next chat appointment with Lumy; (2) collected points in the current and previous week and in total; and (3) personal relaxation goals (number of relaxation exercises) in the current and previous week and outcome goal of the participant.
Theoretical Principles and Operationalization

Overview
The CanRelax app 2.0 implements clinical practice guidelines [35,36]; is grounded in mind-body medicine (MBM) [37,38], the Health Action Process Approach (HAPA) [39], and self-determination theory (SDT) [40]; and includes 39 BCTs (Multimedia Appendix 4 [41]) translated into app features and content. BCTs are active components of behavior change interventions [42] that can influence users’ engagement at both the micro and macro levels. At the micro level, BCTs such as prompts or cues can increase user engagement with the app itself. At the macro level, BCTs can increase engagement with the target behavior (relaxation practice), for example, by using goal setting or self-monitoring features [6,7]. The underlying concept of the intervention flow and the structure of the coaching sessions are informed by generic principles of face-to-face coaching sessions, and we used motivational interviewing (MI) [43,44] aspects as a communication approach. To support the integration of relaxation into everyday routines, we applied the complementing principles of MBM, the HAPA, and SDT as outlined in Figure 4 and detailed in the following sections.

Clinical Practice Guidelines
The CanRelax app 2.0 aims to identify and address distress according to clinical practice guideline recommendations on distress management in people with cancer [35,36] by offering a relaxation and mindfulness intervention specifically designed for individuals with cancer, including initial assessment and monitoring of distress using validated tools such as the Distress Thermometer [29].

MBM Approach
MBM is a resource-oriented approach centered on empowering individuals and supporting healthy, sustainable behaviors [37,38]. Relaxation and mindfulness are widely used self-care interventions in MBM. The CanRelax app 2.0, being a mind-body intervention, provides the opportunity to learn different relaxation techniques along with educational material on distress during cancer, relaxation, and creating healthy habits.

HAPA Framework
Healthy behavior change is at the core of the HAPA. The HAPA focuses on the difficulty of behaving according to one’s intentions and suggests to bridge this intention-behavior gap through perceived self-efficacy, action planning, and coping planning [39]. The CanRelax app 2.0 seeks to enhance self-efficacy and self-management skills through self-enactable BCTs with practical examples of use, such as problem-solving, positive reframing, behavioral experiments, graded tasks, prompts, and self-kindness [41]. Among automatically preselected themes and BCTs (triggered by participants’ interaction with the app), participants can pick the components and topics most relevant to them. The app encourages participants to try new BCTs, determine what works for them, and use these techniques in their daily lives to stay motivated. Participants can also set their own relaxation goals and choose the support they wish to receive from Lumy.

SDT Approach
SDT sees healthy behavior change as closely linked to the satisfaction of basic psychological needs for autonomy, competence, and relatedness [40]. The CanRelax app 2.0 supports these basic needs by offering meaningful rationales and choices, using autonomy-supportive language, acknowledging people’s preferences, recognizing their efforts, and promoting a feeling of being cared for through supportive
coaching sessions and peer support. Peer support is implemented through personalized letters in the app from semifictional people with cancer sharing their struggles and strategies for overcoming obstacles. Personal preferences are acknowledged, for example, by tailoring emojis to participants’ preferred skin tone and providing all chat content in 3 gender options (woman; man; and a gender-neutral option using the gender star, an asterisk placed within German words such as in “Liebe*r Andrea”). Participants select both the skin tone of their emoji and their preferred gender option during the onboarding process. We also let individuals choose their nickname and a form of address they are comfortable with (formal or informal), showing respect for their personal preferences in relation to language use [45].

**MI Approach**

MI is a person-centered communication approach that relates to the selected behavior change theories in that it aims to create a collaborative environment, draws on people’s own goals and values, and supports their autonomy [43,44]. Examples illustrating the integration of MI principles into the app are provided in the Coaching Sessions and Tailoring section.

**Relaxation Exercises**

The app offers 7 different types of relaxation exercises recommended as evidence-based interventions to reduce distress in people with cancer [35,36,46,47]. The relaxation exercises include guided audio recordings of a short meditation (5 minutes), walking meditation (5 minutes), mindfulness meditation (15 minutes), guided imagery (15 minutes), progressive muscle relaxation (15 minutes), body scan (40 minutes), and slow-paced breathing training with visual guidance through gameful visualizations (2-5 minutes; Breeze 2 [48]). The audio files are available in male and female voices with and without background music. The FAQ sidebar submenu in the app provides a selection aid with more information about the different types of relaxation exercises.

**Self-Monitoring Dashboard With Gamification Elements**

The CanRelax app 2.0 tracks relaxation exercises and rewards participants with points as a gamification element. Earned points count toward participants’ self-set weekly relaxation goals. Participants can also earn points by practicing relaxation exercises without the CanRelax app 2.0 (using a different app or without using any app) provided they add this information manually when prompted by Lumy during the coaching sessions. A self-monitoring dashboard illustrates earned points as progress circles. It also provides an overview of the relaxation goals and includes other useful information such as the date and time of the next coaching session.

**Coaching Sessions and Tailoring**

Lumy was developed as a friendly conversational agent that guides participants through the intervention via a series of rule-based, predefined, and personalized conversational turns that simulate the back-and-forth of a real-life conversation. A full coaching session consists of approximately 60 conversational turns (counted in pairs, with one conversational turn consisting of one message from Lumy and one from the participant in response). The conversational flow adapts to the responses chosen by the participants and is enhanced through various ways of tailoring (Textbox 1).

We adopted the structure of a typical face-to-face behavioral coaching session to build the chat sessions in the app [51]. The sessions start with a greeting, followed by small talk about a neutral topic (eg, about the weather) or a “how are you?” sequence and an introduction to the session (including a snooze option to postpone the session). The core part includes assessing the participants’ current state, reviewing previously discussed topics and experiences with BCTs (if applicable), and applying coaching techniques based on MI [43,44]. The implemented techniques focus on building confidence for change (eg, scaling questions, shifting focus away from obstacles and barriers, reframing to offer new and positive interpretations, expressing empathy, affirming, and expressing respect by asking for permission before the conversation starts or before information is shared). After participants have set new relaxation goals, the sessions are summarized to reflect back the main points of the session. An outlook serves as a bridge to the next session, and participants are again encouraged to try out the selected BCTs before the next session (if applicable). The sessions close with the option to adjust the reminder settings and a farewell.
Tailoring concepts and their implementation in the CanRelax app 2.0

- Feedback: Lumy gives feedback on goal setting, goal achievement, and participants’ self-efficacy toward goal achievement. When participants reach their relaxation goals, Lumy celebrates their achievements, and when things do not go well, Lumy tries to offer support.

- Interhuman interaction: in case of urgent need, Lumy encourages interhuman interaction through built-in support to contact relevant services that offer advice and support. Inspired by human coaches, we programmed Lumy to show great attention and commitment, listen with curiosity, reflect, and encourage participants to overcome obstacles. When participants report a challenge they came up against in their practice, they have the option to learn about tips and techniques (behavior change techniques [BCTs]) that can help overcome that challenge. They can choose to skip this section or pick a topic they find interesting among 3 preselected BCTs. Selected BCTs are delivered through personalized in-app letters from semifictional peers, which is another way of supporting interhuman interaction.

- Adaptation: the BCTs are adapted precisely to the reported challenge, and the preselected options are renewed in each coaching session to help keep the sessions interesting.

- User targeting: the concept of user targeting attempts to give participants the impression that the conversation was designed especially for them [49]. We incorporated this concept by identifying participants by their nicknames. We also regard participants’ chosen pronouns (formal or informal), gender identity terms, and emojis as expressions of how participants construct their web identity in the context of the CanRelax app 2.0 [50] and match the chat conversations and the app accordingly.

- Goal setting: goal setting is a BCT that can be used to tailor an intervention and give participants a feeling of progress over time [49]. In CanRelax 2.0, participants’ own weekly relaxation goals and objectives are at the center of the intervention.

- Context awareness: the tailoring concept context awareness aims at providing relevant information considering participants’ (external) situation [49]. We incorporated this by tailoring greeting and farewell messages to the time of day and small talk topics to the season of the year, where applicable.

- Self-learning: CanRelax 2.0 is a self-learning app in the sense that it learns from the interactions with the participants and updates the intervention accordingly. For example, it records the obstacles that participants report and the BCTs they select and uses this information as a bridge to future sessions. To give continuity, the subsequent coaching sessions take up previously discussed topics and include a recap of experiences and learnings (if any) with the new BCTs between sessions.

Iterative Development and Testing

We developed the CanRelax app 2.0 in iterative processes involving input from people with cancer, health professionals, and an interdisciplinary team. The CanRelax app 2.0 builds on a basic app version, which provided relaxation exercises and a reminder function but no other tools or a conversational agent [52]. In version 2.0, we included new features, enhanced functionality, and a solid theory base. During the development process, we conducted usability testing with people with cancer to determine whether they understood and enjoyed the app and whether the app features met their needs. We submitted the usability testing study synopsis to the ethics committee of Zurich, Switzerland, and after review, they stated that the study did not fall under the regulation of the Human Research Act of Switzerland (ethics ID: 2020-00224). A total of 9 individuals with cancer consented to test a prototype of the app, of whom 3 provided detailed feedback, 3 did not test the app in the given time frame, and 3 had technical issues or privacy concerns regarding the test environment. Originally, we planned to conduct the usability tests in person, but due to circumstances related to COVID-19, we had to switch to a fully web-based approach using self-reports. In addition, we thoroughly and repeatedly pretested the app content and features with a multidisciplinary team. The team consisted of professionals with expertise in software engineering, computer science, psychology, psychotherapy, medicine, MBM, nursing, and teaching. Most user feedback was centered on the scripted coaching dialogues with Lumy. We clustered the comments into two main categories and iteratively implemented (1) more variety, in-depth responses, and tailored follow-up questions in the conversation (eg, adjusted the wording of unsatisfactory conversational turns, extended sets of predefined answer options, added links to previously discussed topics, and created unique session openings); and (2) more active choice options with possibilities to skip parts of the conversation, the ability to select topics of personal relevance and interest, the ability to formulate own reminders, and a snooze feature. All improvements were continually refined and tested over 2 years until user satisfaction was achieved.

Assessments

We collected self-reported data (through Lumy and structured in-app questionnaires) and objective app use data at different time points during the 10-week study period. Only the relevant measures considered for this analysis are described in detail in this paper; the measures of the larger study are reported in Multimedia Appendix 1.

Distress and Sociodemographics

At screening and baseline, we collected participants’ self-reported level of distress using a well-known and validated instrument (Distress Thermometer [29]) and sociodemographics such as age, educational attainment, and sex assigned at birth using a structured in-app questionnaire. In the first chat with Lumy, we stored the selected gender identity terms, emoji skin tone modifiers, and preference for formal or informal pronouns (“Du” or “Sie” for “you” in German) to personalize the chat sessions and assessed participants’ initial motivation for downloading the app (outcome goal; 5 forced-choice answer options; see the Intervention section).
Macrolevel Engagement

To answer the research questions of this paper, we combined engagement data on different levels. Data on macrolevel engagement were gathered in the weekly coaching sessions with Lumi. In each session, we asked about relaxation techniques practiced without using the CanRelax app 2.0. The exact wording changed slightly from week to week to help keep the conversation natural (example wording if at least one relaxation exercise was completed in the app: “Did you practice in any other way last week, besides using the CanRelax app?” If no relaxation exercise was completed in the app, the wording was as follows: “Have you practiced in a different way instead, without the CanRelax app?” An example follow-up question if participants answered “yes” would be the following: “In the past seven days, how often have you practiced without using the CanRelax app?”). We assessed reasons for practicing relaxation exercises without the app (if applicable) once per participant and participants’ self-efficacy toward self-set relaxation goals biweekly using a single-item measure developed with the recommended wording for assessing a specific health behavior [53] (“How confident are you that you will reach your relaxation goal next week, even if it gets difficult?”). Participants responded on a visual analog scale implemented as a horizontal slider with values from 0 [not at all confident] to 10 [very confident]).

Microlevel Engagement

At the micro level, we collected participants’ feedback on the app at week 10 with single-choice questions about their favorite feature and the features they would like to change in the app (7 forced-choice answer options in random order) and an option to provide additional information in a free-text field. In addition, the CanRelax app 2.0 tracked the use of different app components (relaxation exercises in the app and coaching sessions with Lumi) over the entire intervention period. Relaxation exercises were considered completed when they were played for 66% of their total run time, and weekly coaching sessions were considered completed when the session closing was reached. We counted the chat sessions 1 to 11 as coaching sessions but not session 0 (onboarding) as completing this session was a requirement for enrollment.

Adherence Definition

We used an adherence definition of at least one relaxation exercise or one coaching session per week for 80% of the weeks during the study period to identify participants who complied fully with the app use suggestions.

Analyses

We conducted descriptive and exploratory data analyses to investigate the data set and thematic analysis of free-text comments. Descriptive statistics were used to report the baseline characteristics of the participants, participants’ self-set goals and self-efficacy, and quantitative in-app feedback. Data visualization methods, supplemented by numerical measures, were used to summarize the main characteristics of the data collected on engagement. We tested for differences in the number of completed relaxation exercises and coaching sessions between prespecified subgroups (distress level at baseline, sex, educational attainment, and age). For this purpose, we conducted a Mann-Whitney U test (in the case of 2 groups) or a Kruskal-Wallis test (for >2 groups) after a detailed investigation of descriptive statistics, checking for outliers using box plots and testing normality using a Shapiro-Wilk normality test and Q-Q plots. Qualitative free-text feedback was analyzed thematically using an inductive approach with the feedback statements as a coding unit, coded into multiple categories where applicable [54]. All analyses were conducted for the entire sample, including those participants who never used the app after onboarding, except for the comparison of relaxation exercises completed using the app versus without using the app. We expected no missing values in baseline variables as completing the questionnaires was a prerequisite for enrollment and participants could not skip questions. Nevertheless, 1 educational attainment response was missing from 1 participant for unknown reasons. Missing values related to the number of exercises or coaching sessions were treated as 0 (no exercise or coaching session completed). Other missing values (educational attainment, self-efficacy, reasons for practicing without the app, and participants’ feedback on the app) were not considered in the analyses.

Statistical analyses and visualizations were conducted using R language (version 4.2.2; R Foundation for Statistical Computing) [55] through RStudio (version 2023.06.0+421; Posit, PBC) [56] using dplyr [57] for data manipulation and summary statistics; ggplot2 [58] for box plots and bar plots; qplot [59] for Q-Q plots; DescTools [60] for median CIs; and the base R stats package to compute the Wilcoxon, Shapiro-Wilk, and Kruskal-Wallis tests.

Ethical Considerations

We submitted the study synopsis to the ethics committee of Zurich, Switzerland, and after review, they stated that the study did not fall under the regulation of the Human Research Act of Switzerland (ethics ID: 2021-01071). The study was conducted according to the Declaration of Helsinki, the Human Research Act, and the Human Research Ordinance. Informed consent was obtained via the app from each participant before enrollment. All data were collected and stored in secure databases and analyzed in a pseudonymized form. Participants did not receive any compensation. Only participants in the intervention group and the nonrandomized third arm received immediate access to the app’s primary features (ie, the relaxation exercises, weekly coaching sessions with Lumi, BCTs, dashboard, reminder notifications, educational video clip, peer support letters, and FAQs), but everyone who downloaded the app had access to a sidebar submenu with useful links (ie, cancer and mental health information leaflets and links to organizations offering support and counseling) and crisis numbers in case urgent help was needed. By using a rule-based conversational agent, we adopted a highly transparent and safe approach compared to artificial intelligence chatbots and had complete control over the content and flow of the coaching sessions [61-63].
Results

Baseline Characteristics

The sample included 77% (77/100) of individuals assigned female at birth and 23% (23/100) of individuals assigned male at birth, and 70% (70/100) self-identified as women, 22% (22/100) self-identified as men, and 8% (8/100) preferred not to disclose their gender. Participants were aged 26 to 79 years (mean 55.6, SD 10.7 years), and 51% (51/100) had a bachelor’s degree or higher. The baseline mean distress level (Distress Thermometer [29]) was 5.6 (SD 2.2), with a mean of 6.9 (SD 1.3) in the high-distress (intervention) group versus 3.1 (SD 0.9) in the low-distress (nonrandomized) group. Baseline characteristics between participants in the high-distress (intervention) group (67/100, 67%) and low-distress (nonrandomized) group (33/100, 33%) were generally comparable except that the high-distress group had fewer participants who installed the app because they were “just curious” (5/67, 7% vs 7/33, 21%). Overall, the most common motivations for installing the app were to improve coping strategies (37/100, 37%) and find inner peace (35/100, 35%; Table 1).

Table 1. Descriptive information about the study sample (N=100).

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>High-distress group (n=67)</th>
<th>Low-distress group (n=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distress level(a), mean (SD)</td>
<td>5.6 (2.2)</td>
<td>6.9 (1.3)</td>
<td>3.1 (0.9)</td>
</tr>
<tr>
<td>Sex assigned at birth, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>77 (77)</td>
<td>53 (79)</td>
<td>24 (73)</td>
</tr>
<tr>
<td>Male</td>
<td>23 (23)</td>
<td>14 (21)</td>
<td>9 (27)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woman</td>
<td>70 (70)</td>
<td>48 (72)</td>
<td>22 (67)</td>
</tr>
<tr>
<td>Man</td>
<td>22 (22)</td>
<td>13 (19)</td>
<td>9 (27)</td>
</tr>
<tr>
<td>Other</td>
<td>8 (8)</td>
<td>6 (9)</td>
<td>2 (6)</td>
</tr>
<tr>
<td>Age (years), n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-44</td>
<td>12 (12)</td>
<td>8 (12)</td>
<td>4 (12)</td>
</tr>
<tr>
<td>45-64</td>
<td>68 (68)</td>
<td>47 (70)</td>
<td>21 (64)</td>
</tr>
<tr>
<td>&gt;64</td>
<td>20 (20)</td>
<td>12 (18)</td>
<td>8 (24)</td>
</tr>
<tr>
<td>Educational attainment, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nontertiary</td>
<td>48 (48)</td>
<td>34 (51)</td>
<td>14 (42)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>51 (51)</td>
<td>32 (48)</td>
<td>19 (58)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Outcome goal, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coping resources</td>
<td>37 (37)</td>
<td>26 (39)</td>
<td>11 (33)</td>
</tr>
<tr>
<td>Inner peace</td>
<td>35 (35)</td>
<td>24 (36)</td>
<td>11 (33)</td>
</tr>
<tr>
<td>Just curious</td>
<td>12 (12)</td>
<td>5 (7)</td>
<td>7 (21)</td>
</tr>
<tr>
<td>Joy in life</td>
<td>10 (10)</td>
<td>8 (12)</td>
<td>2 (6)</td>
</tr>
<tr>
<td>Self-confidence</td>
<td>6 (6)</td>
<td>4 (6)</td>
<td>2 (6)</td>
</tr>
</tbody>
</table>

\(a\)Distress measured using the Distress Thermometer [29] with a rating scale ranging from 0 (no distress) to 10 (extreme distress).

App Engagement

A visual description of the participants’ app use (completed relaxation exercises and coaching sessions) is presented in Multimedia Appendix 5 and Figures 5 and 6, supplemented by the numerical measures in Table 2. During the 10-week study period, 95% (95/100) of the participants used the app at least once after onboarding. These 95 participants completed a total of 4897 relaxation exercises (median 38, IQR 18-73.5) and 714 coaching sessions (median 9, IQR 4-11) over 10 weeks. Of the total number of relaxation exercises, 71.88% (3520/4897) were completed using the CanRelax app 2.0 (95/100, 95% of the participants; median 25.5, IQR 13-55), and 28.12% (1377/4897) were reported as completed without using the app (median 10, IQR 3-19). Among those participants who reported having completed relaxation exercises without using the app, 28% (21/76) specified that they had used different relaxation recordings, 18% (14/76) did not have their smartphones near them, 16% (12/76) knew the exercises by heart, 4% (3/76) preferred to relax without audio recordings, and 34% (26/76) had other reasons for relaxing without using the CanRelax app 2.0.

The proportion of participants who completed at least one relaxation exercise or one coaching session per week (“active app users”) dropped from 88% (88/100) in the first week to...
62% (62/100) in week 10. A total of 64% (64/100) of the participants complied with the app use suggestions per our adherence definition.

Participants’ perceived self-efficacy toward self-set relaxation goals stayed at a median of 8 (0=very low; 10=very high) throughout the 10-week study period, whereas participants raised their relaxation goals. The level of the self-set goals increased from a median of 3 relaxation exercises per week in the first half of the study period (sessions 1 and 3) to a median of 4 exercises per week in the second half (sessions 5, 7, and 9).

App engagement did not vary across prespecified subgroups (ie, distress level at baseline, sex, educational attainment, and age). Mean rank comparisons showed no substantial difference in the number of completed relaxation exercises or coaching sessions among these subgroups (Table 2).

Of the 100 participants, 52 (52%) provided in-app feedback after the 10-week study period (during session 11; Multimedia Appendix 6). A total of 88% (46/52) of the respondents indicated that they “really enjoyed” or “quite enjoyed” chatting with Lumy, and all respondents rated the overall app experience as “very satisfactory” (41/52, 79%) or “quite satisfactory” (11/52, 21%). The favorite app features of the respondents were relaxation exercises (37/52, 71%) and coaching sessions (12/52, 23%). Elements of the app that respondents felt could be improved included “nothing” (29/52, 56%), “something else” than the answer options provided (7/52, 13%), letters from semifictional peers (5/52, 10%), relaxation exercises (4/52, 8%), and in-app questionnaires related to the RCT (4/52, 8%). Of the 52 completed feedback questionnaires, 41 (79%) contained optional free-text comments from participants contextualizing their selected favorite (41 comments) and least favorite (20 comments) app features. Respondents particularly enjoyed the collection of relaxation exercises (12 mentions), liked the format and voices of the exercises (11 mentions), and found that the exercises helped them relax (8 mentions). For example, one respondent stated that the relaxation exercises “are well constructed, with pleasant voices and short.” However, 20% (8/41) of the respondents would have appreciated a wider selection of exercises to choose from. Another main topic that emerged from the analysis was a positive experience of the interaction with Lumy (9 mentions). The coaching sessions were experienced as friendly, uplifting, and encouraging, as seen in the following example:

*It is a very friendly chat with a sense of humor, and it always motivates me.*

Another respondent appreciated “the conscious reflection and looking back. The feeling of being accompanied and encouraged.” However, 5% (2/41) of the respondents also felt that the interaction with Lumy sounded too robotic or was not interactive enough (1 mention each). Tables S1 and S2 in Multimedia Appendix 7 provide an overview of all free-text comments.

*Figure 5.* Comparison of completed relaxation exercises and completed coaching sessions in the high-distress group versus the low-distress group (N=100).
Figure 6. Comparison of completed relaxation exercises and completed coaching sessions in the high-distress group versus the low-distress group for 3 subgroups (N=100).

Table 2. Use of the CanRelax app 2.0 in the first 100 study participants with immediate access to the app, stratified by subgroup (N=100).

<table>
<thead>
<tr>
<th></th>
<th>Participants, n (%)</th>
<th>Completed relaxation exercises over 10 weeks</th>
<th>Completed coaching sessions over 10 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Values, median (IQR)</td>
<td>Values, median (IQR)</td>
</tr>
<tr>
<td><strong>Entire sample</strong></td>
<td>100 (100)</td>
<td>34.5 (14-70.75)</td>
<td>8 (4-11)</td>
</tr>
<tr>
<td><strong>Subgroup</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distress</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>67 (67)</td>
<td>41 (15.5-77)</td>
<td>9 (4-11)</td>
</tr>
<tr>
<td>Low</td>
<td>33 (33)</td>
<td>30 (8-61)</td>
<td>8 (2-11)</td>
</tr>
<tr>
<td><strong>Sex assigned at birth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>77 (77)</td>
<td>40 (15-70)</td>
<td>9 (3-11)</td>
</tr>
<tr>
<td>Male</td>
<td>23 (23)</td>
<td>30 (8.5-64.5)</td>
<td>7 (4.5-11)</td>
</tr>
<tr>
<td><strong>Educational attainment</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nontertiary</td>
<td>48 (48)</td>
<td>31.5 (12.75-64.75)</td>
<td>8 (2.75-11)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>51 (51)</td>
<td>41 (18-78)</td>
<td>9 (3-11)</td>
</tr>
<tr>
<td><strong>Age group (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-44</td>
<td>11 (11)</td>
<td>27 (12.5-39.5)</td>
<td>9 (3.5-10)</td>
</tr>
<tr>
<td>45-64</td>
<td>68 (68)</td>
<td>34.5 (13.25-73.25)</td>
<td>8 (3-11)</td>
</tr>
<tr>
<td>&gt;64</td>
<td>20 (20)</td>
<td>52.5 (19.25-72.25)</td>
<td>11 (6.75-11)</td>
</tr>
</tbody>
</table>

<sup>a</sup> 2-sided <i>P</i> values derived from the Mann-Whitney <i>U</i> test (distress, sex, and educational attainment) and Kruskal-Wallis test (age group).

<sup>b</sup> Not applicable.

<sup>c</sup> 1 missing value.

**Discussion**

**Principal Findings**

Overall, engagement with the CanRelax app 2.0 declined over the study period but stayed relatively high, with 62% (62/100) of participants actively using the app in week 10. Engagement was unrelated to participant characteristics such as level of distress at baseline, sex assigned at birth, educational attainment, or age. More than one-quarter (1377/4897, 28.12%) of the relaxation exercises were completed without using the app, supporting the need for assessing engagement on a macro level.
Participants’ self-efficacy remained stable at a high level. At the same time, participants raised their relaxation goals, which indicates a potential relative increase in self-efficacy. Participants who completed the intervention highly valued the app. Free-text comments suggested that a wider variety of relaxation exercises would further enhance the user experience.

Comparison With Prior Work

Engagement rates with cancer-related digital interventions tend to be higher than in other populations, but high variability in engagement measures and intervention components and lack of a threshold for acceptable engagement make it difficult to compare findings across studies. Reviews of empirical studies using cancer-related digital interventions have reported use rates between 70% and 100% [64,65]. These high use rates contrast with the generally low engagement with mHealth apps reported for individuals with other health conditions [2,10,11,13-15] and suggest that people with cancer might be particularly inclined to improve their health and change certain health behaviors through mHealth apps. Stressful life events such as the diagnosis and treatment of cancer potentially serve as catalysts for behavior change [66,67]. Nonetheless, comparing engagement across studies is difficult as there are no standards regarding the assessment, reporting, and interpretation of engagement with mHealth apps. In a recent review, every primary study stated that their apps achieved good engagement despite large differences in criteria used to assess engagement and a range of reported engagement rates from 35% to 100% [16]. This shows an urgent need for standards for assessing, reporting, and interpreting engagement with mHealth apps [16].

Fully automated mHealth studies with no human support are prone to low engagement rates, but there is great potential for increasing engagement using fully automated social components, behavior change theory, and design principles of successful commercial apps. Most mHealth apps in research settings provide human support, whereas popular commercial apps are typically unguided. Human support is known to positively influence engagement and effectiveness but drastically limits the scalability of mHealth apps [3,18]. Despite this limitation, most mHealth apps in research settings provide human support at varying levels—from high support through guided interventions (ie, involving guidance from a trained professional, eg, through live videoconferencing or web-based workshops) to lower levels of support through study processes (eg, screening visits or telephone surveys conducted by the study team). In the rare studies available on unguided cancer-related mHealth apps with no human support [68-70], engagement rates were <50%. One reason could be that existing researcher-developed apps are not engaging enough and, therefore, need human support to motivate participants [18,71]. This may be less the case for popular commercial apps, which are typically unguided (eg, Headspace and Calm) [18]. Thus, there is great potential for unguided research apps to improve user engagement and the generalizability of research findings to real-life settings if they learn from successful commercial apps. An example of an mHealth study with a low level of human support is the CanRelax 1.0 feasibility study [52]. The CanRelax app 1.0 was a fully automated mHealth app, but study processes such as enrollment were supported by study staff. The authors classified 54% (54/100) of participants as continuous app users in week 10 [52]. In comparison, engagement with the enhanced CanRelax app 2.0 in week 10 improved to 62% (62/100) even though we provided no human support and used stricter definitions of engagement. It is possible that the fully automated social components in the CanRelax app 2.0, such as the weekly coaching sessions with Lumy, compensated for the lack of human support. This aligns with recent research underscoring the potential of conversational agents to positively impact engagement with mHealth apps [3,72-77]. We demonstrated this potential by combining a conversational agent with a theoretical foundation and incorporating key design principles inspired by highly engaging commercial apps (eg, inclusive avatar and visuals).

Existing findings on the impact of participant characteristics on engagement are inconsistent [78]. In our analyses, engagement was not associated with the demographics (sex assigned at birth, educational attainment, and age) or psychological characteristics (level of distress) of the participants. These results contradict the findings of earlier studies that showed higher engagement in female individuals [6,52], individuals with higher educational attainment [6,15], younger [15] or older individuals [6], and individuals with higher baseline distress [15,52]. In the CanRelax app 2.0, the content and design features implemented to increase engagement might have succeeded in reaching those groups of people who needed a little extra encouragement and possibly helped level out differences in engagement among subgroups. Given the inconsistencies in the literature, identifying participant characteristics and other factors that influence engagement is an exciting topic for future studies.

Our findings support the feasibility and value of assessing macrolevel engagement in mHealth behavior change interventions. Although the conceptualization of engagement as a multifaceted construct is widely accepted, macrolevel engagement is rarely assessed in mHealth app studies. We approached this gap by examining engagement on multiple levels and showed considerable engagement with the target behavior (ie, relaxation) beyond app use. First, nearly one-third of all completed relaxation exercises (1377/4897, 28.12%) were practiced without using the CanRelax app 2.0. Relaxation techniques can be practiced in different ways depending on one’s experiences, needs, and preferences; for example, beginners could start with guided relaxation via audio recordings (or in-person sessions) and later move on to more silent, self-guided relaxation exercises. In our study, examining only those exercises practiced using the app would have given an incomplete and potentially misleading picture of participants’ engagement with relaxation practices. Second, median self-efficacy remained high even as relaxation goals increased, indicating that participants felt encouraged to tackle challenging tasks and were engaged in the behavior change process [24].

Data on macrolevel engagement are necessary to understand how engagement with an mHealth app changes over time and how these engagement patterns relate to the intended health outcomes. Baglione et al [17] found that high baseline distress was associated with initially higher engagement that declined over time, whereas the engagement of the group of participants with lower baseline distress increased over the course of a...
7-week intervention, resulting in similar engagement levels in both groups at week 7. Siebenhüner et al [79] examined the associations between distress and adherence (ie, app use) in the CanRelax app 1.0 and showed that a decrease in the level of distress over time (ie, an improvement in health outcomes) was associated with lower adherence. However, the authors did not assess engagement with the target behavior in daily life. Without this information, it remains unclear whether participants with improved distress stopped using the app because they disengaged from the intervention or no longer needed the app’s support to continue the new behavior [5]. As lower app use could be associated with higher engagement at the macro level, the suggested “adherence benefit paradox” [79] might not be a paradox after all but could even be considered the goal of a successful mHealth app [2].

Limitations

Our study is subject to common sources of bias that can affect the internal validity and generalizability of the findings. One potential source of bias is the use of self-reported data. To mitigate potential self-reporting bias, we combined self-reported and objectively tracked data in the assessment of engagement. Feedback was only collected from participants who completed the coaching session with Lumy in week 11. As it is possible that only those who enjoyed the app completed this session, feedback might be positively biased. Another potential source of bias is selection bias as our study focused on a group of highly motivated participants. Initial motivation for study participation was needed as participants had no contact with the research team but self-downloaded the app and self-enrolled in the study if they fulfilled the inclusion criteria. Selection bias is also indicated by female individuals being overrepresented in our sample. To improve the generalizability of our study, we used broad recruitment strategies and successfully recruited participants with lower than tertiary education. We also abstained from using research strategies to increase motivation and engagement (eg, compensation for study participation) that would differ from usual real-world app use settings. Another limitation is that we did not consider past engagement with relaxation in our analyses. Participants could have already established a regular relaxation practice before the study; still, engagement with a new app is not necessarily linked to previous experience with relaxation. A third limitation is due to technical issues with the CanRelax app 2.0 during the study, which could have reduced engagement. For example, we did not provide an easy solution to transfer the CanRelax app 2.0 to a new smartphone. Participants with new smartphones had to reach out for technical support and usually had to wait several weeks until they could continue to use the app where they left off. To avoid this problem, individuals must create an account in the future.

Clinical Implications

For a positive impact on health outcomes on a large scale, mHealth apps need to be scalable, engaging to users, and effective. Scalability is a great advantage of fully automated mHealth apps, but these apps tend to suffer from low engagement rates threatening their effectiveness. Our findings show that successful engagement can be achieved with fully automated mHealth apps that are highly tailored, include fully automated social components and BCTs based on theory and evidence, and are developed with design principles used by popular commercial apps. These results provide a valuable context for subsequent outcome evaluations and add to research on optimizing fully automated digital health interventions.

Conclusions

The CanRelax app 2.0 achieved similar engagement to that of other cancer-related mHealth apps even though we used stricter criteria for engagement than other studies and provided no human support. The implemented theory- and evidence-based design principles and fully automated social components, such as a conversational agent that simulated human support, might have compensated for the lack of human involvement and contributed to enhanced engagement at both a micro and a macro level. Our findings underline that engagement is a complex and multifaceted construct and that measures at the macro level are particularly valuable to assess engagement not only with the app itself but also with the larger target behavior, which is ultimately, the goal of an mHealth app.

Acknowledgments

The authors thank and acknowledge all study participants for their time and responses and all usability test participants for contributing throughout the app development process. The authors thank all organizations and individuals who contributed to facilitating or conducting recruitment. The authors thank Anita Thomae (University Hospital Zurich) for creating the app’s educational video and participating in usability testing. The authors thank Manuela Oehler (University Hospital Zurich) for designing the visual representation of Lumy and the project website and flyers and for participating in multiple rounds of usability testing. The authors thank Isabelle Wyder (University of Zurich) for participating in multiple rounds of usability testing. This work was supported by the Swiss Cancer Research foundation (KFS 4556-08-2018).

Authors’ Contributions

JB, CMW, ME, TK, and SS conceptualized the study. SS and JB conceptualized the intervention. SS created the intervention, designed the conversational flow, wrote the scripts for the chatbot coaching sessions and the letters from semifictional peers, and implemented the app on the MobileCoach platform. JB reviewed and contributed to the intervention content. CMW provided audio recordings of relaxation exercises. FS and PS set up the MobileCoach infrastructure, developed project-specific features, deployed the app, and provided technical support during the study. SS, FS, PS, JB, ME, and TK were involved in app testing. SS
drafted the manuscript and conducted the analyses. JB helped interpret the results and contributed to reviewing and finalizing the manuscript. All authors reviewed and contributed to the manuscript and approved the final version.

**Conflicts of Interest**

ME received institutional research grants from Kaiku Health, reports grants from Bristol Myers Squibb and Roche, and institutional fees as a Scientific Advisory. ME is also a Board Member and Consultant from Roche, outside the submitted work. CMW has active research grants to the university for digital health projects from the DIZH, the Swiss Cancer Research foundation, the German health care Innovation Fund, and New senselab GmbH. CMW also received honoraria from Swiss hospitals for scientific presentations on digitalization and AI in medicine and integrative oncology. TK, FS, and PS are developers and promoters of the open-source software platform MobileCoach. TK, FS, and PS are affiliated with the Centre for Digital Health Interventions, a joint initiative of the Institute for Implementation Science in Health Care, University of Zurich; the Department of Management, Technology, and Economics at ETH Zurich; and the Institute of Technology Management and the School of Medicine at the University of St. Gallen. The Centre for Digital Health Interventions is funded in part by CSS, a Swiss health insurer, Movie Next, an Austrian health insurer, and MTIP, a Swiss digital health investor. TK was also a co-founder of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways. However, neither CSS, Movie Next, MTIP, nor Pathmate Technologies was involved in this study. JB received honoraria for workshops on digital health. The remaining authors have no conflicts of interest to declare.

**Multimedia Appendix 1**
CanRelax randomized controlled trial study design and assessments.

[PDF File (Adobe PDF File), 227 KB - cancer_v10i1e52386_app1.pdf ]

**Multimedia Appendix 2**
CanRelax coaching structure.

[PNG File, 22 KB - cancer_v10i1e52386_app2.png ]

**Multimedia Appendix 3**
CanRelax avatar icon.

[PNG File, 11 KB - cancer_v10i1e52386_app3.png ]

**Multimedia Appendix 4**
Behavior change techniques in the CanRelax app 2.0.

[PDF File (Adobe PDF File), 173 KB - cancer_v10i1e52386_app4.pdf ]

**Multimedia Appendix 5**
Active app users over 10 weeks (N=100).

[PNG File, 7 KB - cancer_v10i1e52386_app5.png ]

**Multimedia Appendix 6**
CanRelax in-app feedback.

[PNG File, 22 KB - cancer_v10i1e52386_app6.png ]

**Multimedia Appendix 7**
CanRelax free-text feedback.

[PDF File (Adobe PDF File), 153 KB - cancer_v10i1e52386_app7.pdf ]

**References**


Lukic YX, Teepe GW, Fleisch E, Kowatsch T. Breathing as an input modality in a gameful breathing training app (Breeze 2): development and evaluation study. JMIR Serious Games 2022 Aug 16;10(3):e9186 [FREE Full text] [doi: 10.2196/39186] [Medline: 35972793]


URL: https://www.r-project.org/ [accessed 2024-04-29]


Abbreviations

BCT: behavior change technique
FAQ: frequently asked question
HAPA: Health Action Process Approach
MBM: mind-body medicine
mHealth: mobile health
MI: motivational interviewing
RCT: randomized controlled trial
SDT: self-determination theory

©Sonja Schläpfer, Fabian Schneider, Prabhakaran Santhanam, Manuela Eicher, Tobias Kowatsch, Claudia M Witt, Jürgen Barth. Originally published in JMIR Cancer (https://cancer.jmir.org), 31.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Heart Rate Monitoring Among Breast Cancer Survivors: Quantitative Study of Device Agreement in a Community-Based Exercise Program

Lindsey L Page1, MPH, PhD; Jason Fanning2, PhD; Connor Phipps1, PhD; Ann Berger3, APRN-CNS, AOCNS, FAAN, PhD; Elizabeth Reed4, MD; Diane Ehlers1,5, PhD

1Department of Neurological Sciences, University of Nebraska Medical Center, Omaha, NE, United States
2Department of Health and Exercise Science, Wake Forest University, Winston-Salem, NC, United States
3College of Nursing, University of Nebraska Medical Center, Omaha, NE, United States
4Department of Internal Medicine, University of Nebraska Medical Center, Omaha, NE, United States
5Department of Quantitative Health Sciences, Division of Epidemiology, Mayo Clinic Arizona, Rochester, AZ, United States

Corresponding Author:
Diane Ehlers, PhD
Department of Quantitative Health Sciences
Division of Epidemiology
Mayo Clinic Arizona
200 First Street SW
Rochester, AZ, 55905
United States
Phone: 1 408 574 2739
Email: ehlers.diane@mayo.edu

Abstract

Background: Exercise intensity (e.g., target heart rate [HR]) is a fundamental component of exercise prescription to elicit health benefits in cancer survivors. Despite the validity of chest-worn monitors, their feasibility in community and unsupervised exercise settings may be challenging. As wearable technology continues to improve, consumer-based wearable sensors may represent an accessible alternative to traditional monitoring, offering additional advantages.

Objective: The purpose of this study was to examine the agreement between the Polar H10 chest monitor and Fitbit Inspire HR for HR measurement in breast cancer survivors enrolled in the intervention arm of a randomized, pilot exercise trial.

Methods: Participants included breast cancer survivors (N=14; aged 38-72 years) randomized to a 12-week aerobic exercise program. This program consisted of three 60-minute, moderate-intensity walking sessions per week, either in small groups or one-on-one, facilitated by a certified exercise physiologist and held at local community fitness centers. As originally designed, the exercise prescription included 36 supervised sessions at a fitness center. However, due to the COVID-19 pandemic, the number of supervised sessions varied depending on whether participants enrolled before or after March 2020. During each exercise session, HR (in beats per minute) was concurrently measured via a Polar H10 chest monitor and a wrist-worn Fitbit Inspire HR at 5 stages: pre-exercise rest; midpoint of warm-up; midpoint of exercise session; midpoint of cool-down; and postexercise recovery. The exercise physiologist recorded the participant’s HR from each device at the midpoint of each stage. HR agreement between the Polar H10 monitor and the Fitbit was assessed using Lin concordance correlation coefficient (rc) with a 95% CI. Lin rc ranges from 0 to 1.00, with 0 indicating no concordance and 1.00 indicating perfect concordance. Relative error rates were calculated to examine differences across exercise session stages.

Results: Data were available for 200 supervised sessions across the sample (session per participant: mean 13.33, SD 13.7). By exercise session stage, agreement between the Polar H10 monitor and the Fitbit was highest during pre-exercise seated rest (rc=0.76, 95% CI 0.70-0.81) and postexercise seated recovery (rc=0.89, 95% CI 0.86-0.92), followed by the midpoint of exercise (rc=0.63, 95% CI 0.55-0.70) and cool-down (rc=0.68, 95% CI 0.60-0.74). The agreement was lowest during warm-up (rc=0.39, 95% CI 0.27-0.49). Relative error rates ranged from −3.91% to 3.09%, with 0 indicating no concordance and 1.00 indicating perfect concordance. Relative error rates were calculated to examine differences across exercise session stages.
Conclusions: The Fitbit overestimated HR during peak exercise intensity, posing risks for overexercising, which may not be safe for breast cancer survivors’ fitness levels. While the Fitbit Inspire HR may be used to estimate exercise HR, precautions are needed when considering participant safety and data interpretation.

Trial Registration: Clinicaltrials.gov NCT03980626; https://clinicaltrials.gov/study/NCT03980626?term=NCT03980626&rank=1

KEYWORDS
wearable devices; exercise prescription; validity; photoplethysmography; monitoring; wearables; devices; exercise; heart rate; breast cancer; cancer; cancer survivor; community; chest monitor; Fitbit; recovery; safety

Introduction

Wearable sensors have gained traction in both commercial and research sectors [1], with a projected 156 million units to be purchased in 2024 [2]. These devices use microelectronic triaxial accelerometers to measure steps, energy expenditure, sleep, and time spent in different intensities of activity and photoplethysmography above the wrist to measure heart rate (HR). These data, along with options for goal setting, can be used to help individuals self-monitor and increase their daily physical activity (PA) [3]. The ease and utility of these devices have led to their adoption in health promotion research for continuous measurement of health behaviors and as behavior change tools [4].

Traditional research-grade monitors are costly, lack consumer-friendly designs, provide little opportunity for user interaction with the device, often evaluate only 1 dimension of daily activities (eg, HR, motion, or sleep only), and have limited, real-time data transfer capacity [3,5,6]. In contrast, consumer-grade monitors continuously collect and transfer data through Bluetooth and web-based platforms to allow for data collection across months or even years [3]. Additionally, participant data can be easily monitored and accessed via web-based platforms at any point during the data collection period. The increasing number of peer-reviewed publications and National Institutes of Health–funded grant proposals, which include consumer-grade, wrist-worn monitors, emphasizes the utility of these devices in research settings [1,7].

Many of these devices now measure HR, a key component of aerobic exercise prescription. Although electrocardiogram (ECG) is widely accepted as the gold standard for assessing HR during exercise, chest-worn monitors also have well-documented validity for measuring HR [8]. However, like ECG, they may be inconvenient or prohibitive in community-based and unsupervised exercise settings due to necessary receivers and participant discomfort. In contrast, newer devices are increasingly being designed for wear on the forearm or wrist. Commercially available wearable sensors, such as the Fitbit, represent an accessible, multifunctional alternative to HR monitoring in exercise. Appropriate exercise intensity, often expressed as a percentage of HR reserve, is a fundamental dimension of exercise prescription for achieving the health benefits of exercise [9]. For example, cancer survivors begin to reduce fatigue symptoms with a minimum dose of aerobic activity at 45% of their HR reserve, whereas benefits for other symptoms (ie, anxiety, depression, and physical function) begin at a minimum dose of 60% of their HR reserve [10]. To improve dissemination and uptake of exercise prescriptions in clinical or community-based settings, it is critical that survivors have user-friendly methods to independently monitor exercise prescription components.

There is an increasing number of exercise oncology studies that use commercial wearable sensors to intervene in PA behaviors and reduce cancer-related symptom burden, particularly among breast cancer survivors, and evidence indicates that wearable sensors are effective, feasible, and user-friendly for breast cancer survivors in exercise interventions [11-13]. Although these studies have helped bolster the utility of wearable sensors in PA promotion research, they have failed to provide any detail on intensity or HR monitoring during their respective interventions. Many exercise interventions in cancer populations are adopting community-based, hybrid, and unsupervised designs [14]. Therefore, it is integral that researchers understand the capacity, and limitations, of commercially available wearable sensors in providing accurate measurements of HR to monitor participants’ safety and compliance with the exercise prescription.

In the general population, the reliability of popular, commercially available activity and HR monitors has been previously examined with varying agreement between commercial products and traditional ECG monitoring [8,15-17]. Unfortunately, many of these data have been collected in controlled settings with predetermined treadmill speeds in young, healthy adult participants. The dearth of literature assessing commercially available, wrist-worn HR monitors in clinical populations during training sessions limits the utility of these devices in less controlled environments. Therefore, the purpose of this study was to examine the agreement between a commercially available, wrist-worn wearable sensor (Fitbit Inspire HR; Fitbit Inc) and a traditional chest-worn monitor (Polar H10; Polar Electro OY) for HR measurement in breast cancer survivors at different stages of exercise in a community-based program. It is hypothesized that HR monitor agreement in this study will be highest at periods of pre- and postexercise rest and lowest during the exercise session when participants were exercising at higher intensities.

Methods

Study Design and Participants

The Study on Physical Activity’s Relationship with Cancer and Cognition (SPARCC) was a randomized exercise trial in which 30 women diagnosed with breast cancer were randomized to a 12-week moderate-intensity aerobic exercise program (n=15).
or usual care (n=15). Our study includes only those women randomized to the exercise group with valid Fitbit and Polar HR data (n=14), as exercise HR was not collected from women in the usual care control group.

Eligibility criteria for this study included the following: female participants aged 21 years or older; postmenopausal at the time of diagnosis; first, primary diagnosis of breast cancer (stage I-IIIa); within 3-24 months of completing surgery, chemotherapy, or radiation therapy; self-reported an average of <60 minutes of moderate to vigorous PA per week for the previous 6 months; having received physician’s clearance to participate in an exercise program; and randomized to the 12-week aerobic exercise program in the SPARCC study. Participants were recruited from a midwestern academic medical center, a private cancer center, and the community (eg, via flyers to community organizations, social media posts, and word of mouth). Interested individuals were scheduled for a phone appointment to confirm eligibility, absence of neurological disorders, and interest in participating in the study. Eligible women were then asked to attend an in-person or Zoom-based orientation session to receive more information about the study, decide if they would like to participate, sign the Institutional Review Board (IRB)-approved informed consent, and schedule baseline testing appointments. After baseline data collection was complete, participants were randomized in a block design to the 12-week aerobic exercise program or usual care.

**Table 1. Exercise prescription.**

<table>
<thead>
<tr>
<th>Week</th>
<th>Intensity (rating of perceived exertion), range</th>
<th>Intensity (% heart rate reserve), range</th>
<th>Duration (minutes), range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9-11</td>
<td>45-50</td>
<td>15-20</td>
</tr>
<tr>
<td>2</td>
<td>9-11</td>
<td>45-50</td>
<td>20-25</td>
</tr>
<tr>
<td>3</td>
<td>9-11</td>
<td>45-50</td>
<td>25-30</td>
</tr>
<tr>
<td>4</td>
<td>11-13</td>
<td>50-55</td>
<td>25-30</td>
</tr>
<tr>
<td>5</td>
<td>11-13</td>
<td>50-55</td>
<td>30-35</td>
</tr>
<tr>
<td>6</td>
<td>11-13</td>
<td>50-55</td>
<td>35-45</td>
</tr>
<tr>
<td>7</td>
<td>13-15</td>
<td>55-65</td>
<td>35-45</td>
</tr>
<tr>
<td>8</td>
<td>13-15</td>
<td>55-65</td>
<td>40-50</td>
</tr>
<tr>
<td>9</td>
<td>13-15</td>
<td>55-65</td>
<td>40-50</td>
</tr>
<tr>
<td>10</td>
<td>15-17</td>
<td>65-75</td>
<td>40-50</td>
</tr>
<tr>
<td>11</td>
<td>15-17</td>
<td>65-75</td>
<td>45-50</td>
</tr>
<tr>
<td>12</td>
<td>15-17</td>
<td>65-75</td>
<td>45-50</td>
</tr>
</tbody>
</table>

As originally designed, all 36 exercise sessions were scheduled to be delivered by the exercise physiologist in the supervised, community-based setting. However, due to the COVID-19 pandemic, the exercise program was modified for some participants to include both supervised and unsupervised sessions. Participants who were in the middle of the intervention in March 2020, were transitioned to unsupervised exercise with weekly Zoom-based counseling from their trainer. Participants enrolled after March 2020, engaged in only 4 supervised exercise sessions held once per week in the research team’s exercise laboratory in weeks 1-4. All sessions in weeks 7-12 were unsupervised, home-based sessions with weekly Zoom-based exercise counseling. Across the study, 4 breast cancer survivors completed 36 supervised sessions as originally designed, 5 were in the middle of the intervention in March 2020, and 6 were enrolled after March 2020 and received 4 supervised sessions. Depending on the enrollment time (ie, before or after the COVID-19 public health restrictions), participants engaged in an average of 13.33 (SD 13.71) supervised sessions (range: 4-36). Participants who received the original intervention received their fitness center membership during the study. Those enrolled during or after the onset of the COVID-19 pandemic were offered a fitness center membership when it was safe to do so based on local IRB and public health considerations.

**Ethical Considerations**

This study was approved by The University of Nebraska Medical Center IRB and is registered with the National Institutes of Health’s ClinicalTrials.gov (NCT03980626). All participants provided written informed consent prior to participation. All data presented herein were deidentified using study identification numbers and stored separately from participants, identifiers. Data were collected and managed using applications hosted by the study institution (ie, Research Electronic Data Capture [REDCap] or Box Enterprise) [18,19]. Participants did not receive payment for their participation in this research but received a Fitbit Inspire HR that was theirs to keep. All participants were offered a 3-month membership to a local fitness center.

**Exercise Protocol**

Breast cancer survivors randomized to the exercise program engaged in small group or one-on-one, moderate-intensity walking sessions facilitated by a certified exercise physiologist. These sessions were held at local community fitness centers 3 times per week for 1 hour per session. All participants completed a treadmill-based submaximal cardiopulmonary exercise test prior to randomization to establish baseline fitness and safety and inform individualized exercise prescriptions (Table 1).
requirements. Only supervised sessions (n=200) were included in our analysis.

The exercise program was progressive in nature such that the volume of exercise increased across weeks from 15-20 minutes of walking in weeks 1-2 to 40-45 minutes in weeks 8-12 and from 40%-55% estimated HR reserve in week 1 to 65%-70% HR reserve in weeks 9-12 (Table 1). All sessions began with a 5-minute light-intensity walking warm-up and ended with an active cool-down including light walking and static stretches (Figure 1). The exercise program was designed to follow American College of Sports Medicine guidelines for exercise in cancer survivors [10].

**Figure 1.** Exercise session stages for heart rate measurement.

**Measures**

**HR Monitors**

All participants received a Fitbit Inspire HR sensor to wear on their nondominant wrist for the duration of the study and a Polar H10 chest strap to wear during supervised exercise. The Fitbit was chosen for this study because it is one of the most popular wrist-worn activity trackers, represents approximately 20% of the commercial wearable sensor market, and has sold 63 million devices worldwide in the last decade [3]. The Fitbit Inspire HR measures HR via optical photoplethysmography, which is processed using proprietary algorithms. Briefly, this is done by shining a light on the skin, assessing the reflected light, using algorithms to determine changes in blood volume based upon reflected light, and calculating HR based on oscillations in blood volume [20,21]. The Polar H10 chest strap monitor was chosen as the comparator device because it has high validity with ECG, the gold standard for measuring exercise HR [8]. Exercise trainers fit participants with the Polar H10 monitors, placed on the distal sternum, at the start of each supervised exercise session and used Polar HR readings to adjust treadmill speed and grade to meet prescribed exercise intensity. HR was measured concurrently using the Fitbit and Polar monitors at 5 stages of the exercise session: pre-exercise seated rest; midpoint of the 5-minute warm-up; midpoint of the moderate-intensity exercise session; midpoint of the 5-minute cool-down; and after a 5-minute seated recovery (Figure 1).

**Demographic and Clinical Information**

Participant demographics (ie, age, race, education, income, employment status, marital status, and comorbid conditions) were self-reported and collected via REDCap hosted by the study institution. Clinical information on breast cancer diagnosis and treatment were obtained via electronic medical records. BMI was calculated from height and weight measured via the Seca 703 scale and stadiometer (Seca Corp) by the study staff at a baseline testing visit.

**Data Analysis**

HR from the Polar monitor was operationalized as the criterion measure and used to assess absolute and relative paired differences between monitors [8]. Agreement between HR measurements was assessed using Lin concordance correlation coefficient ($r_c$) with 95% CIs. This test measures the degree to which the paired observations fall on the identity line and defines statistical agreement as $r_c \geq 0.85$ [22]. The agreement was also represented visually across stages using Bland-Altman plots with upper and lower limits set using 95% CIs [23,24]. Absolute paired differences were calculated by subtracting the Fitbit-measured HR from the Polar-measured HR at each stage of the exercise session. Relative paired differences were calculated as relative error rate (RER) across exercise session stages [25,26], as follows:

$$RER = \frac{(\text{Polar HR measurement} - \text{Fitbit HR measurement}) \times 100}{\text{Polar HR measurement}}$$

Negative resultant RERs are indicative of an overestimation of HR by the Fitbit, and positive RERs are indicative of an underestimation of HR by the Fitbit, as compared to the Polar monitor. Data were analyzed using SPSS (version 27; IBM Corp) and RStudio (version 1.3.1093; R Core Team).

**Results**

**Participant Characteristics**

Participants (mean age 63.1, SD 8.7 years) were White women with a history of early-stage breast cancer; on average, overweight; and physically inactive (Table 2). Additionally, more than 1 quarter of participants had a history of clinically diagnosed anxiety or depression at the time of enrollment. Breast cancer survivors in this study were enrolled approximately 17 months after their diagnosis. Participants’ breast cancer treatments included surgery, radiation, and chemotherapy; however, most women in this study did not receive chemotherapy. One participant randomized to the exercise program did not have valid Polar data for supervised sessions and was, therefore, excluded from the analysis (N=14).
Table 2. Participant characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>63.07 (8.66)</td>
</tr>
<tr>
<td>Non-Hispanic White, n (%)</td>
<td>13 (93)</td>
</tr>
<tr>
<td>Bachelor’s degree, n (%)</td>
<td>8 (57)</td>
</tr>
<tr>
<td>Income &gt;US $40,000 per year, n (%)</td>
<td>8 (57)</td>
</tr>
<tr>
<td>Employed full-time, n (%)</td>
<td>8 (57)</td>
</tr>
<tr>
<td>Married, n (%)</td>
<td>12 (86)</td>
</tr>
<tr>
<td>Comorbidities$^a$, mean (SD)</td>
<td>2.38 (2.10)</td>
</tr>
<tr>
<td>β-Blocker medication use, n (%)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Antihypertensive medication use, n (%)</td>
<td>1 (7)</td>
</tr>
<tr>
<td>Diagnosed with depression, n (%)</td>
<td>4 (29)</td>
</tr>
<tr>
<td>Diagnosed with anxiety, n (%)</td>
<td>4 (29)</td>
</tr>
<tr>
<td><strong>Clinical features</strong></td>
<td></td>
</tr>
<tr>
<td>Cancer stage, n (%)</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>12 (86)</td>
</tr>
<tr>
<td>II</td>
<td>2 (14)</td>
</tr>
<tr>
<td>Time since diagnosis (months), mean (SD)</td>
<td>16.57 (7.97)</td>
</tr>
<tr>
<td>Chemotherapy, n (%)</td>
<td>3 (21)</td>
</tr>
<tr>
<td>Radiation, n (%)</td>
<td>9 (64)</td>
</tr>
<tr>
<td>Hormonal therapy, n (%)</td>
<td>9 (64)</td>
</tr>
<tr>
<td>Months of hormonal therapy, mean (SD)</td>
<td>13.44 (5.64)</td>
</tr>
</tbody>
</table>

$^a$Comorbid conditions include diagnosed arthritis, osteoporosis, asthma, chronic obstructive pulmonary disease, angina, heart failure, previous myocardial infarction, vascular disease, diabetes, tremors, gastrointestinal disease, visual impairment, hearing impairment, degenerative disk disease, anxiety, and depression.

**HR Monitor Agreement**

Agreement between the Fitbit and Polar HR monitors was highest during seated rest at postexercise ($r=0.89$, 95% CI 0.86-0.92) and pre-exercise ($r=0.76$, 95% CI 0.70-0.81). This was followed by the midpoint of the moderate-intensity exercise session ($r=0.63$, 95% CI 0.55-0.70) and cool-down ($r=0.68$, 95% CI 0.60-0.74). The warm-up was associated with the lowest level of agreement between monitors (0.39, 95% CI 0.27-0.49). RERs ranged from −3.91% to 3.09% and were most pronounced during warm-up (RER: mean −3.91%, SD 11.92%). When inaccurate, the Fitbit overestimated HR during most stages of the exercise session (RER range: −3.91% to −0.52%), except at the midpoint of moderate-intensity exercise, where HR was underestimated (RER 3.09%). RERs and concordance coefficients are provided in Table 3, and Bland-Altman plots are provided in Figure 2.
Table 3. Heart rate (beats per minute) monitor differences according to the stage of the exercise.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Heart rate, mean (SD)</th>
<th>Fitbit differences from Polar H10, mean (SD)</th>
<th>Agreement ($r_c$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Polar H10</td>
<td>Fitbit</td>
<td>Paired difference</td>
</tr>
<tr>
<td>Pre-exercise rest</td>
<td>78.00 (7.68)</td>
<td>78.61 (8.74)</td>
<td>−1.34 (5.53)</td>
</tr>
<tr>
<td>Warm-up</td>
<td>94.78 (7.91)</td>
<td>96.26 (9.17)</td>
<td>−2.96 (10.27)</td>
</tr>
<tr>
<td>Exercise</td>
<td>117.04 (8.10)</td>
<td>115.15 (9.3)</td>
<td>3.77 (6.98)</td>
</tr>
<tr>
<td>Cool-down</td>
<td>103.56 (6.68)</td>
<td>104.24 (7.90)</td>
<td>−0.72 (5.83)</td>
</tr>
<tr>
<td>Postexercise rest</td>
<td>85.86 (7.42)</td>
<td>85.99 (7.71)</td>
<td>−0.38 (3.53)</td>
</tr>
</tbody>
</table>

$^a$Relative error rate.

$^b$Lin concordance correlation coefficient.

Figure 2. Bland-Altman plots. Bland-Altman plots depict the average heart rate (Polar H10 and Fitbit) by the relative difference between the two measures for each session by stage in the exercise protocol. Points on the plots indicate individual sessions, solid lines indicate the mean difference across the sample, and dashed lines indicate upper and lower bounds for each stage.
Discussion

Principal Results

This study was the first to examine the agreement between the wrist-worn Fitbit Inspire HR and chest-worn Polar H10 HR monitors in cancer survivors and during an exercise intervention. Major findings from this study indicate that the Fitbit monitor did not produce statistically accurate measures of HR during most exercise stages, especially the warm-up stage; however, agreement during seated rest (pre- and postexercise), midpoint of exercise, and cool-down were associated with moderate agreement between devices. Data also suggest that the Fitbit underestimated HR only during the primary, aerobic portion of the exercise session (ie, the midpoint of the exercise session), which may have serious implications for safety and compliance monitoring in exercise programs. This study extends the current literature on consumer-grade, wrist-worn HR monitors and provides data to inform future studies hoping to use consumer-grade sensors to monitor safety, exercise program compliance, and longitudinal behavioral patterns in cancer survivors [8,25,27]. This is particularly important as exercise interventions become less centralized and hybrid and unsupervised approaches increase in prevalence [28-30].

Comparison With Prior Work

The wrist-worn, Fitbit Inspire HR monitor accurately measured HR only during seated rest postexercise compared to the chest-worn Polar H10 monitor. Although pre-exercise seated rest, midpoint of moderate-intensity exercise, and cool-down also exhibited high levels of agreement, they did not reach statistical agreement as defined by Lin concordance correlation coefficient [22]. Results from previous studies have found that Fitbit devices are most accurate in measuring HR during low-intensity activities where the wrist is moving in a repetitive fashion [17]. Nevertheless, in contrast to these previous studies, warm-up represented the period of the poorest agreement. It is also unclear why the Fitbit accurately measured HR during postbut not pre-exercise seated rest; however, both pre-and postexercise seated rest reflected the highest levels of agreement with the Polar monitor, consistent with previous findings [8,16].

Although HR was highest during the midpoint of the exercise session and cool-down, these stages represented higher levels of agreement when compared to the warm-up stage. This may, in part, be due to the slower speeds at which breast cancer survivors were walking in this study, as compared to healthy, young, or middle-aged adults in other studies [8,16,17,31]. While previous research found that lower treadmill speeds showed the highest agreement, overall speeds in those studies ranged from 2 to 9 miles per hour [16,17]. In comparison, breast cancer survivors in this study did not reach speeds greater than 3.5 miles per hour. It is possible that speeds in this study were more similar to light-intensity walking in previous studies, which would align more closely with the results from this study [16,17]. This does not, however, explain the poor agreement during the warm-up and cool-down stages.

Of note, the RER at the midpoint of the exercise session indicates that the Fitbit monitor underestimated HR as compared to the Polar chest strap. Previous studies have reported similar trends in Fitbit data as compared to traditional ECG monitoring [8,16,17,31]. One study found the Fitbit Ionic to be comparable to other wrist-worn monitors and statistically accurate at rest [17], while another found that the Fitbit Blaze provided the least accurate optically measured HR [8]. In contrast to the findings presented here, these studies also found that higher-intensity activity led to decreased accuracy in HR measurement [8,17]. However, differences in Fitbit accuracy at peak exercise intensity between previous studies and data presented in this study may be due to the lower absolute intensity in this study, as both previous studies were conducted among athletes [8,17]. Given the generally lower intensity of exercise prescribed to cancer survivors, participants may not reach an exercise intensity high enough for devices to decrease in accuracy during steady-state exercise. This should, theoretically, reinforce the utility of Fitbits in cancer survivor populations.

Despite this, underestimation of exercise HR may be problematic in programs using Fitbit to monitor intensity during exercise in cancer survivors for several reasons. First, participants may be asked to increase the intensity of a session to achieve the prescribed HR range. If the Fitbit monitor underestimates HR, as it did in this study, participants who reach the prescribed HR as measured by Fitbit may be exercising at an intensity higher than that prescribed, leading to concerns regarding participant safety—particularly if the session is unsupervised. In a previous analysis of exercise prescription adherence in this sample, data indicated that participants only met prescribed intensity during supervised sessions 57.5% of the time when assessed via Fitbit HR, as compared to 92.2% when measured via Polar. However, adherence to the prescribed intensity via Fitbit was 83.2% during unsupervised sessions after the onset of the COVID-19 pandemic [32]. These data, when taken together with the findings of this study, suggest that participants may have been exercising above their prescribed HR range during unsupervised sessions. Although many breast cancer survivors may comfortably exercise at higher intensities, reliance upon consumer-grade wearable sensors only may introduce safety concerns not previously observed in more traditional, controlled exercise trials. Additionally, future studies that use Fitbit to measure the dose of exercise required to effect specific outcomes (eg, cancer-related fatigue and cognitive performance) may underestimate the required intensity of activity to elicit an effect. Although these devices may have utility in exercise oncology, it is critical that researchers and practitioners are aware of limitations that may increase the risk of adverse events or decrease methodological rigor in quantifying compliance with exercise prescriptions.

Limitations

Although this study is one of the first to investigate HR agreement between the Fitbit Inspire HR and Polar H10 chest monitor during community-based exercise in breast cancer survivors, it is not without limitations. First, this study was performed on a small sample of breast cancer survivors. Our sample was primarily comprised of White, educated women with early-stage breast cancer, which may not be representative of the larger breast cancer population. For example, women in this study were also not on any β-Blockers, and only 1 participant reported antihypertensive medication use. Although
this improves the internal validity of this study due to the lack of HR suppression, it is likely not representative of the broader breast cancer survivor population in the United States [33]. Future studies would be strengthened by the inclusion of a larger, more diverse sample of breast cancer survivors with more supervised exercise sessions.

Additionally, exercise physiologists were available to help and provide feedback on using the devices during the exercise sessions, making it unclear to what extent user error would influence Fitbit’s accuracy in unsupervised exercise settings. Fitbit HR measurements in this study were also compared to Polar chest strap monitors, rather than the gold standard ECG. This may have introduced systematic error in evaluating agreement. Finally, the total number of sessions observed in this study was fewer than originally planned (ie, 200 observed vs 540 planned) due to COVID-19 required adaptations. It is unclear whether additional observations would have changed or stabilized results relative to device agreement.

Conclusions
Overall, Fitbit devices with HR monitoring capabilities may be useful for participant monitoring in exercise oncology studies. Researchers should use caution when using these devices, however, as they likely do not provide accurate HR measurement during critical stages of exercise sessions. This study showed that Fitbit monitors were only statistically accurate during seated rest and likely underestimated HR during steady-state exercise. As a result, Fitbit HR measurements are likely best for estimating exercise intensity rather than evaluating compliance with exercise prescriptions. Because of their ease and potential utility in behavioral PA interventions, future studies should further examine the agreement between wrist-worn wearable sensors and a gold-standard measurement of HR, such as ECG, in a larger, more representative sample of breast cancer survivors. Additionally, studies should analyze agreement by relative HR intensity to determine whether Fitbit may be more appropriate for specific exercise prescriptions.

Acknowledgments
Funding for this study was provided by the University of Nebraska Medical Center Fred and Pamela Buffett Cancer Center Support Grant (P30CA036727) as well as the Nebraska Tobacco Settlement Biomedical Research Development Fund (NTSBRDF). The funders were not involved in the study design, implementation, or dissemination activities. The authors would like to acknowledge Tiffany Dudley, BS, American Collect of Sports Medicine (ACSM) Certified Exercise Physiologist (ACSM-EP), and Joseph Scharfenkamp, MS, ACSM Clinical Exercise Physiologist (ACSM-CEP) for the implementation of the exercise intervention and their flexibility in pivoting from in-person to remote delivery in March 2020.

Data Availability
The data sets generated during this study are available from the corresponding author on reasonable request.

Authors' Contributions
LLP carried out conceptualization, data curation, formal analysis, investigation, methodology, project administration, and writing of the original draft. JP was in charge of data curation, methodology, and software, as well as reviewing and editing of the manuscript. CP contributed to the formal analysis and visualization as well as reviewing and editing of the manuscript; AMB and ECR carried out the methodology and contributed to the reviewing and editing of the manuscript. ECR also provided resources for the study. DKE was in charge of the conceptualization, funding acquisition, methodology, project administration, resources, and supervision, as well as reviewing and editing of the manuscript.

Conflicts of Interest
None declared.

References


11. Hartman SJ, Nelson SH, Weiner LS. Patterns of Fitbit Use and Activity LevelsPatterns of Fitbit use and activity levels throughout a physical activity intervention: exploratory analysis from a randomized controlled trials Throughout a Physical Activity Intervention: Exploratory Analysis from a Randomized Controlled Trial. JMIIR Mhealth Uhealth 2018 Feb 05;6(2):e29 [FREE Full text] [doi: 10.2196/mhealth.8503] [Medline: 29402761]


Abbreviations

ECG: electrocardiogram
HR: heart rate
IRB: Institutional review board
PA: physical activity
REDCap: Research Electronic Data Capture
RER: relative error rate
SPARCC: Study on Physical Activity’s Relationship with Cancer and Cognition
Do Measures of Real-World Physical Behavior Provide Insights Into the Well-Being and Physical Function of Cancer Survivors? Cross-Sectional Analysis

Shelby L Bachman¹, PhD; Emma Gomes², BS; Suvekshya Aryal¹, MPH; David Cellá³, PhD; Ieuan Clay¹, PhD; Kate Lyden¹, PhD; Heather J Leach², PhD

¹VivoSense, Inc, Newport Coast, CA, United States
²Department of Health and Exercise Science, Colorado State University, Fort Collins, CO, United States
³Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:
Shelby L Bachman, PhD
VivoSense, Inc
27 Dorian
Newport Coast, CA, 92657
United States
Phone: 1 8588768486
Email: shelby.bachman@vivosense.com

Abstract

Background: As the number of cancer survivors increases, maintaining health-related quality of life in cancer survivorship is a priority. This necessitates accurate and reliable methods to assess how cancer survivors are feeling and functioning. Real-world digital measures derived from wearable sensors offer potential for monitoring well-being and physical function in cancer survivorship, but questions surrounding the clinical utility of these measures remain to be answered.

Objective: In this secondary analysis, we used 2 existing data sets to examine how measures of real-world physical behavior, captured with a wearable accelerometer, were related to aerobic fitness and self-reported well-being and physical function in a sample of individuals who had completed cancer treatment.

Methods: Overall, 86 disease-free cancer survivors aged 21-85 years completed self-report assessments of well-being and physical function, as well as a submaximal exercise test that was used to estimate their aerobic fitness, quantified as predicted submaximal oxygen uptake (VO₂). A thigh-worn accelerometer was used to monitor participants’ real-world physical behavior for 7 days. Accelerometry data were used to calculate average values of the following measures of physical behavior: sedentary time, step counts, time in light and moderate to vigorous physical activity, time and weighted median cadence in stepping bouts over 1 minute, and peak 30-second cadence.

Results: Spearman correlation analyses indicated that 6 (86%) of the 7 accelerometry-derived measures of real-world physical behavior were not significantly correlated with Functional Assessment of Cancer Therapy-General total well-being or linked Patient-Reported Outcomes Measurement Information System-Physical Function scores (Ps≥.08). In contrast, all but one of the physical behavior measures were significantly correlated with submaximal VO₂ (Ps≤.03). Comparing these associations using likelihood ratio tests, we found that step counts, time in stepping bouts over 1 minute, and time in moderate to vigorous physical activity were more strongly associated with submaximal VO₂ than with self-reported well-being or physical function (Ps≤.03). In contrast, cadence in stepping bouts over 1 minute and peak 30-second cadence were not more associated with submaximal VO₂ than with the self-reported measures (Ps≥.08).

Conclusions: In a sample of disease-free cancer survivors, we found that several measures of real-world physical behavior were more associated with aerobic fitness than with self-reported well-being and physical function. These results highlight the possibility that in individuals who have completed cancer treatment, measures of real-world physical behavior may provide additional information compared with self-reported and performance measures. To advance the appropriate use of digital measures in oncology clinical research, further research evaluating the clinical utility of real-world physical behavior over time in large, representative samples of cancer survivors is warranted.

Trial Registration: ClinicalTrials.gov NCT03781154; https://clinicaltrials.gov/ct2/show/NCT03781154
Introduction

Background

As a result of progress in early cancer detection and the development of effective anticancer therapies, the number of individuals who have survived cancer is increasing. As of 2022, >18 million individuals in the United States were living with a history of cancer [1]. In the future, this number is projected to increase as the aging population grows and cancer screening, treatment, and survivorship care continue to advance [2,3]. Although increases in cancer survivorship are cause for optimism, clinicians and regulators alike are increasingly interested in ensuring that increases in cancer survival rates translate to additional years of good quality life [4,5].

Cancer and its treatments have major impacts on health-related quality of life [6]. These effects persist long into survivorship, with more than one-third of cancer survivors reporting that symptoms persist after treatment ends [7-9]. Across studies, individuals off cancer treatment, henceforward referred to as cancer survivors, report reductions in physical performance, fatigue, sleep problems, mood disturbances, and pain as long-term symptoms, even years after being disease free [9-11]. The impacts of cancer and its treatments are also associated with poorer outcomes and survival in the long term. For instance, individuals who experience a greater health burden from cancer symptoms are at an elevated risk of developing chronic comorbidities [12]. Furthermore, among adults with a history of cancer, both depression [13,14] and reduced physical function [15-18] are associated with an increased risk of mortality after controlling for confounding variables. At the same time, there is accumulating evidence that in cancer survivorship, health-promoting behaviors have positive impacts; for instance, exercise interventions have been demonstrated to improve health-related quality of life, objectively assessed physical function, and aerobic fitness in cancer survivors [19-21].

Therefore, understanding and considering the long-term impacts of anticancer therapies on health-related quality of life should be an integral component of assessing risk-benefit profiles during both regulatory and medical decision-making. This necessitates methods to accurately and reliably capture features of health-related quality of life that are important to cancer survivors. Established methods to assess these constructs in oncology clinical research include patient-reported assessments of global and domain-specific well-being, clinician-reported assessments of functional capacity, and performance assessments that capture physical performance capacity [22,23]. Collectively, these assessments have a range of limitations: patient-reported outcome assessments are burdensome and prone to floor and ceiling effects [24], clinician-reported outcome assessments exhibit limited interobserver reliability [25,26], and performance outcome assessments do not reflect many of the day-to-day functional challenges experienced by those with a history of cancer. Together, these limitations raise the question of how to best capture how cancer survivors are feeling and functioning in their real-world environments.

In the midst of a digital transformation in medicine, there is a growing interest in digital health technologies as measurement tools in oncology clinical care and research [27,28]. In particular, wearable sensors such as accelerometers have the potential to address some of the limitations of the established assessments of health-related quality of life in oncology [29]. These technologies can capture aspects of everyday physical behavior remotely (in individuals’ lived environment), passively (as individuals go about their daily lives), and continuously (with high granularity) [28,30]. These devices can furthermore capture many domains of physical behavior, including aspects of gait, mobility, posture, physical activity, and sedentary behavior [31-33]. Alongside established outcome assessments, these measures may provide rich insights into the real-world well-being and physical function in cancer survivorship [34-36].

The use of wearable sensors as monitoring tools in oncology clinical research is on the rise [37], but despite their potential for capturing how individuals feel and function in their real-world environments, these tools have not been widely adopted for assessing treatment efficacy or monitoring in cancer clinical research. Furthermore, across trials that have deployed wearable sensors, there is little standardization regarding which outcome measures are included, as well as the definitions of those measures [37]. Together with this lack of standardization, a potential reason for the limited adoption of digital measures derived from wearable sensors is that there is limited clinical validation evidence linking specific digital measures of real-world physical behavior to gold-standard outcome measures commonly used in oncology clinical research (ie, patient-reported, clinician-reported, and performance outcomes) [29].

Objectives

In this secondary analysis, we aimed to gain insight into how various digital measures of real-world physical behavior, captured with wearable sensors, can provide an additional understanding of health-related quality of life following cancer treatment. To do so, we leveraged data from 2 previous studies of individuals who had completed cancer treatment to test whether an array of digital measures of real-world physical behavior, measured with a wearable accelerometer over a 1-week period, were related to self-reported and performance measures of physical function. First, we examined associations between real-world physical behavior and self-reported well-being and physical function. Next, we examined how real-world physical behavior was related to aerobic fitness, captured with a submaximal exercise test performed in the clinic. Finally, we compared these patterns of associations to determine whether real-world physical behavior was more closely related
to self-reported well-being and physical function or to aerobic fitness.

Methods

Overview

Data were collected as part of 2 studies. Study 1 was a cross-sectional study conducted at Colorado State University between January 2020 and June 2021 and aimed to examine how reallocating time to physical activity affected body composition and quality of life in individuals who had completed cancer treatment [38,39]. Study 2 was a randomized clinical trial conducted at the University of Colorado Anschutz Medical Campus and Colorado State University and aimed to examine the feasibility and preliminary effects of a videoconference physical activity intervention in individuals who had completed treatment for colorectal cancer [40,41]. For study 2, only data collected at the baseline measurement time point (ie, before the initiation of the intervention) were used. These data were collected between February 2021 and July 2022. For increased statistical power, we combined data from studies 1 and 2.

Ethical Considerations

Study 1

The study protocol was approved by Colorado State University Institutional Review Board (IRB #19-8914H). All participants provided written, informed consent before participation and were compensated US $25 for participation. When providing consent, participants consented to their deidentified data being used for future studies. Data were deidentified before analyses.

Study 2

The study protocol was approved by the University of Colorado Institutional Review Board (IRB #18-2436). Informed consent was obtained from all participants. As a part of this process, participants consented to their deidentified data being used for research purposes beyond the primary study aims. Participants were compensated up to US $75 for participation. Data were deidentified before analyses.

Participants

Study 1

Participants in study 1 were recruited from local and regional cancer centers and the Colorado State University Center for Healthy Aging using flyers, presentations, and email postings. Eligible participants were aged ≥18 years at the time of their cancer diagnosis and within 60 months of treatment completion at the time of study participation.

Study 2

Participants in study 2 were recruited from the University of Colorado Cancer Center, survivor support organizations, and community outreach using mailed letters, flyers, and social media platforms. Eligible participants (1) were fluent in English, (2) had access to a computer or phone with internet and a camera, (3) stated willingness to comply with all study procedures and be available for the duration of the study, (4) were male or female individuals aged ≥40 years at the time of diagnosis, (5) had histologically confirmed cancer of the colon or rectum (stages II–IV) if treated with curative intent, completed resection or other surgery 3 to 60 months before enrollment, received chemotherapy and/or radiation therapy within the previous year with at least 1 cycle of intended chemotherapy completed (not necessary to have completed all cycles), and had no plans for additional chemotherapy or radiation therapy. Exclusion criteria were evidence of metastatic disease, existing participation in at least 150 minutes per week of at least moderate intensity physical activity, being pregnant or planning to become pregnant, and known contraindications for exercise.

Procedure

We aimed to test relationships between participants’ real-world physical behavior, self-reported well-being and physical function, and aerobic fitness; therefore, we focused only on relevant assessments that were included in both studies. These assessments are described in subsequent sections.

Assessments of Self-Reported Well-Being and Physical Function

In the laboratory, participants completed a series of questionnaires in which they reported demographic information and information about their cancer diagnosis and types of treatment completed. They also completed the Functional Assessment of Cancer Therapy-General (FACT-G), a 27-item instrument designed to assess health-related quality of life in individuals with cancer along 4 dimensions: physical, functional, emotional, and social well-being [42]. For FACT-G and its subscales, higher scores indicate better well-being.

Assessment of Aerobic Fitness

Following the questionnaires, participants completed a submaximal exercise test that involved a modified Balke Treadmill Test. The modified Balke Treadmill protocol consisted of a 3-minute warm-up at a treadmill speed of 2.5 mph. Following the warm-up, participants entered stage 1 of the test at 0% grade and 2.5 mph. Every 3 minutes, participants entered a new stage, increasing the treadmill grade by 2.5% until 70% heart rate reserve was reached or until there was a safety indication to stop the exercise test. Heart rate was collected every minute throughout the protocol. A measure of aerobic fitness, that is, predicted oxygen uptake (VO2) at 70% heart rate reserve, was then calculated according to the following formula (for women [43]; for men [44]), where T denotes the test duration (ie, time to reach 70% heart rate reserve):

\[
\text{Predicted submaximal VO}_2 = \frac{1.38 (T) + 5.22}{(1)} \\
\text{Predicted submaximal VO}_2 = \frac{1.44 (T) + 14.99}{(2)}
\]

Assessment of Real-World Physical Behavior

At the end of the laboratory visit, participants were instructed that during the subsequent 7-day period, their real-world behavior would be monitored continuously using an activPAL3 activity monitor (PAL Technologies Ltd), worn on the thigh [45]. Using an accelerometer to sense limb position and activity,
activPAL can discriminate between the activities of lying, sitting, standing, and stepping and therefore allows for the calculation of time spent in various physical activity categories [46-48]. The sensor identifies reciprocal leg movements as steps, and based on the detected steps, measures including cadence and time in stepping bouts of various durations can be calculated [49].

Participants were each given an activPAL and instructed regarding proper use and wear of the device. Each participant was instructed to wear the device on their thigh for 7 days in their real-world environments. A 7-day monitoring period has been demonstrated to provide sufficient accelerometer data for generating reliable estimates of various measures of real-world physical behavior [50-52]. After the remote monitoring period, participants returned their devices to the laboratory. If their appointment to return the device was >7 days after the beginning of the remote monitoring period, participants were permitted to wear the device longer than 7 days to avoid losing it. All available activPAL data were used for analysis.

Analysis

**Linkage of PROMIS-Physical Function Scores**

To assess self-reported physical function, we first calculated scores on a 5-item subset of the FACT-G physical well-being subscale. The 5 items in the subset were “I have a lack of energy,” “Because of my physical condition, I have trouble meeting the needs of my family,” “I have pain,” “I feel ill,” and “I am forced to spend time in bed.” This 5-item subset excluded 2 items on the FACT-G physical well-being subscale: “I have nausea” and “I am bothered by the side effects of treatment.” These 5-item subset scores were linked to T scores on a custom subset of the Patient-Reported Outcomes Measurement Information System-Physical Function (PROMIS-PF) calibrated item bank, using an established linkage method [53]. PROMIS-PF is a tool for assessing physical function in oncology clinical research [54,55], for which higher T scores indicate better physical function. We used the linkage procedure described by Kaat et al [53] rather than administering the PROMIS-PF assessment directly.

**Summarization of Self-Reported and Performance Measures in the Sample**

For the purposes of analysis, FACT-G scores, linked PROMIS-PF T scores, and submaximal VO₂ values were treated as continuous variables. Summary statistics were used to summarize the sample in terms of FACT-G total well-being scores, FACT-G physical well-being subscale scores, scores on the 5-item subset of the FACT-G physical well-being subscale used for linkage to PROMIS-PF, linked PROMIS-PF T scores, and submaximal VO₂. Ceiling effects, defined as the percentage of the sample achieving the maximum possible score [56], were calculated for each self-reported measure. The skewness and kurtosis of each self-reported measure’s distribution were also calculated.

**Calculation of Measures of Real-World Physical Behavior**

Average daily nonwear time, defined as the time in which participants did not wear the activPAL monitor, was calculated for each participant. A valid day was considered as the one in which a participant wore the monitor for at least 10 hours; only participants with at least 4 valid days during the remote monitoring period were included for analysis [57].

The activPAL proprietary software, PALbatch (version 8.11.1.63; PAL Technologies), was used to access summaries of recorded data and whole recording outcomes from the real-world monitoring period. Measures of interest included average daily time spent sedentary (ie, secondary lying, defined as sitting or lying not classified as primary lying); time in light physical activity; time in moderate to vigorous physical activity; and step count. Average daily time in light and moderate to vigorous intensity activity was calculated using established approaches [47]. In addition, we calculated the average daily time that each participant spent in stepping bouts of ≥1 minute in duration. Finally, we extracted 2 measures of cadence: weighted median cadence in stepping bouts of ≥1 minute across all valid days, as well as the number of steps in any 30-second recording period (“peak 30 s cadence”) across all valid days, a measure that is thought to reflect an individual’s best natural effort [58-60]. Summary statistics were used to characterize the sample in terms of the various measures of real-world physical behavior.

**Intercorrelations Among Related Measures**

As preliminary tests for expected intercorrelations among the self-reported measures and among the measures of real-world physical behavior, we performed Spearman correlation analyses.

**Associations With Measures of Real-World Physical Behavior**

Pairwise Spearman correlation analyses were then used to test for associations between each measure of real-world physical behavior and (1) the self-reported measures and (2) aerobic fitness. These analyses were repeated in a partial Spearman correlation framework to account for the effects of age, sex, BMI, time since diagnosis, and cancer stage at diagnosis on each association. In addition, to test for differences in physical behavior based on the level of self-reported physical function and well-being, we first performed a tertile split of each self-reported measure and a median split of aerobic fitness. A median split instead of a tertile split was performed for aerobic fitness since fewer participants had values of submaximal VO₂ available compared with the self-reported measures. In cases where scores were equal to a tertile value, they were assigned such that the resulting splits reflecting high, medium, and low scores were approximately equal in size. Then, we used 2-tailed pairwise Welch t tests and Mann-Whitney U tests to compare the splits in terms of the various measures of real-world physical behavior. Welch t tests were used to compare splits in terms of measures that did not exhibit deviations from normality, whereas Mann-Whitney U tests were performed to compare splits in terms of measures that exhibited deviations from normality. Deviation from normality was indicated by a statistically
significant Shapiro-Wilk test result. For each comparison of the splits of self-reported measures, $P$ values were adjusted for multiple comparisons using Holm method [61]. For comprehensiveness, we also performed Spearman correlation analyses to test for associations between aerobic fitness and each of the self-reported measures.

**Comparison of Associations With Measures of Real-World Physical Behavior**

A series of likelihood ratio tests was used to determine if the strength of associations with real-world physical behavior differed between the self-reported measures and aerobic fitness. The following steps were performed for each measure of real-world physical behavior. Here, we describe the process for FACT-G total well-being, but the same process was used for FACT-G physical well-being, FACT-G physical well-being 5-item subset, and linked PROMIS-PF T scores:

1. One multiple linear regression model was fit, with all measures (FACT-G total well-being, submaximal VO$_2$, age, sex, BMI, time since diagnosis, and cancer stage) regressed onto the measure of real-world physical behavior.
2. A second multiple linear regression model was fit, which was identical to the first, with the exception that the regression coefficients for FACT-G total well-being and submaximal VO$_2$ were constrained to equality.
3. A likelihood ratio test was performed to compare the fits of the first (unconstrained) and second (constrained) models; a significant test result indicated that constraining the coefficients to equality led to a significantly poorer model fit.

**Exploratory Analysis of Associations With Activity Fragmentation**

For additional insights into real-world physical behavior, we calculated measures of activity fragmentation, reflecting how participants accumulated their total activity and sedentary time across the days of the remote monitoring period [62]. More fragmented activity patterns have been associated with increased mortality risk, reduced physical function as measured with in-clinic physical performance tests, and fatigability [63,64]. Using a similar approach as mentioned in the Comparison of Associations With Measures of Real-World Physical Behavior section, we tested whether the various measures of activity fragmentation were more associated with the self-reported measures or with aerobic fitness (Multimedia Appendix 1 [62,65,66]).

For each analysis comparing regression coefficients, data were restricted to include only those participants with no missing values for the respective measures being compared (ie, the self-reported measure of interest and submaximal VO$_2$). In addition, for all linear regression analyses, continuous variables were standardized, and binary variables were coded with a sum contrast coding scheme before analysis. All analyses were performed with R (version 4.1.2; The R Foundation).

### Results

#### Overview

For study 1, we screened 101 individuals for participation and enrolled 59 (58.4%) individuals; 2 enrolled participants did not undergo remote monitoring. For study 2, we screened 1149 individuals and enrolled 29 (2.52%; screening details for study 2 are described fully in the study by Leach et al [41]). A total of 86 participants across both studies (study 1: n=57, 66%; study 2: n=29, 34%) completed remote monitoring of physical behavior and were included in the combined data set for analysis. Characteristics of participants included in the combined data set are summarized in Table 1. A comparison of participants in the 2 study samples in terms of demographics, cancer diagnosis, and treatment information is provided in Table S1 in Multimedia Appendix 1. Across studies 1 and 2, the most common cancer types at diagnosis were breast (n=21, 24%), colon (n=20, 23%), and colorectal cancers (n=13, 15%). Detailed information on the distribution of cancer types across the 2 studies is presented in Table S2 in Multimedia Appendix 1.

All participants who underwent remote monitoring had valid activPAL data for least 4 days during the remote monitoring period. Participants had an average of 7.2 (SD 1.4; range 4-13) days of valid data and an average of 35.6 (SD 46.1; range 0-177.8) minutes of nonwear time per day. One participant in study 1 did not complete the FACT-G physical well-being subscale. Due to some in-person assessments being suspended during the COVID-19 pandemic, submaximal VO$_2$ values were only available for 37% (21/57) of the participants in study 1. Submaximal VO$_2$ values were available for all but 1 participant (28/29, 97%) in study 2 (due to an equipment malfunction). This yielded a total of 49 participants across both studies with available values for submaximal VO$_2$.

A summary of the measures of self-reported well-being and physical function, aerobic fitness, and real-world physical behavior is presented in Table 2. Although no ceiling effects were observed for FACT-G total well-being scores, moderate ceiling effects were observed for the FACT-G physical well-being subscale, the FACT-G physical well-being 5-item subset, and linked PROMIS-PF T scores, with 19% (16/85), 22% (19/85), and 22% (19/85) of the participants having the maximum score, respectively. Three of the self-reported measures had skewness <-1; distributions of all measures are visualized in Figures S1 and S2 in Multimedia Appendix 1.

As expected, FACT-G total well-being, FACT-G physical well-being, FACT-G physical well-being 5-item subset, and linked PROMIS-PF T scores were significantly correlated (Figure S3 in Multimedia Appendix 1). Similarly, the various measures of real-world physical behavior exhibited mostly expected intercorrelations (Figure S4 in Multimedia Appendix 1).
Table 1. Participant characteristics (n=86).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean (SD; range)</td>
<td>55.4 (12.9); 21-85</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>61 (71)</td>
</tr>
<tr>
<td>Male</td>
<td>25 (29)</td>
</tr>
<tr>
<td>BMI, mean (SD; range)</td>
<td>27.4 (5.2; 18-43)</td>
</tr>
<tr>
<td>Education level, n (%)</td>
<td></td>
</tr>
<tr>
<td>12th grade or less</td>
<td>0 (0)</td>
</tr>
<tr>
<td>High school graduate or GED</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Some college, AA degree, or technical school</td>
<td>21 (24)</td>
</tr>
<tr>
<td>College graduate (Bachelor’s)</td>
<td>29 (34)</td>
</tr>
<tr>
<td>Graduate degree (masters or doctorate)</td>
<td>32 (37)</td>
</tr>
<tr>
<td>Prefer not to answer</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Time since diagnosis (mo), mean (SD; range)</td>
<td>32 (25.5; 2-211)</td>
</tr>
<tr>
<td>Time since last treatment (mo), mean (SD; range)</td>
<td>21.2 (17.0; 0-60)</td>
</tr>
<tr>
<td>Cancer stage at diagnosis, n (%)</td>
<td></td>
</tr>
<tr>
<td>0(^a)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>I</td>
<td>15 (17)</td>
</tr>
<tr>
<td>II</td>
<td>22 (26)</td>
</tr>
<tr>
<td>III</td>
<td>29 (34)</td>
</tr>
<tr>
<td>IV</td>
<td>9 (11)</td>
</tr>
<tr>
<td>Unsure</td>
<td>7 (8)</td>
</tr>
<tr>
<td>Cancer treatment</td>
<td></td>
</tr>
<tr>
<td>Had any treatment</td>
<td>86 (100)</td>
</tr>
<tr>
<td>Had chemotherapy</td>
<td>65 (76)</td>
</tr>
<tr>
<td>Had radiation</td>
<td>42 (49)</td>
</tr>
<tr>
<td>Had surgery</td>
<td>76 (88)</td>
</tr>
<tr>
<td>Had other</td>
<td>12 (14)</td>
</tr>
<tr>
<td>Number of treatment types, n (%)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>13 (15)</td>
</tr>
<tr>
<td>2</td>
<td>41 (47.7)</td>
</tr>
<tr>
<td>3</td>
<td>28 (32.6)</td>
</tr>
<tr>
<td>4</td>
<td>4 (4.7)</td>
</tr>
</tbody>
</table>

\(^a\)Stage 0 indicates evidence of abnormal cells in situ.
Table 2. Summary of measures of self-reported well-being and physical function, aerobic fitness, and real-world physical behavior.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Values, mean (SD; range)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-reported well-being and physical function</strong></td>
<td></td>
</tr>
<tr>
<td>FACT-G total well-being (0-108)</td>
<td>87.9 (13.3; 41-107)</td>
</tr>
<tr>
<td>FACT-G physical well-being subscale (0-28)</td>
<td>24.4 (3.7; 12-28)</td>
</tr>
<tr>
<td>FACT-G physical well-being subscale 5-item subset (0-20)</td>
<td>17.4 (2.6; 9-20)</td>
</tr>
<tr>
<td>Linked PROMIS-PF Physical Function T-score (19-61)</td>
<td>51.1 (7.1; 35-61)</td>
</tr>
<tr>
<td><strong>Aerobic fitness</strong></td>
<td></td>
</tr>
<tr>
<td>Predicted submaximal VO(_2) (mL/kg/min)</td>
<td>29.1 (9.9; 10.0-50.0)</td>
</tr>
<tr>
<td><strong>Real-world physical behavior</strong></td>
<td></td>
</tr>
<tr>
<td>Daily sedentary time (min)</td>
<td>582.1 (102.8; 295.1-819.4)</td>
</tr>
<tr>
<td>Daily step count</td>
<td>6916.3 (2704.5; 1413-17,501)</td>
</tr>
<tr>
<td>Daily time in light activity (min)</td>
<td>305.1 (97.2; 103.8-551.3)</td>
</tr>
<tr>
<td>Daily time in moderate to vigorous activity (min)</td>
<td>4.0 (6.6; 0.0-58.2)</td>
</tr>
<tr>
<td>Daily time in stepping bouts ≥1 min (min)</td>
<td>25.3 (19.7; 0.2-107.7)</td>
</tr>
<tr>
<td>Weighted median cadence in stepping bouts ≥1 min (steps/min)</td>
<td>98.7 (12.1; 56.5-126.2)</td>
</tr>
<tr>
<td>Peak 30-second cadence (steps/min)</td>
<td>67.2 (8.6; 42.0-86.0)</td>
</tr>
</tbody>
</table>

\(^a\)FACT-G: Functional Assessment of Cancer Therapy-General.  
\(^b\)PROMIS: Patient-Reported Outcomes Measurement Information System.  
\(^c\)VO\(_2\): submaximal oxygen uptake.

Most Measures of Real-World Physical Behavior Were Not Associated With Self-Reported Well-Being or Physical Function

Spearman correlations with real-world physical behavior are depicted in Figure 1. The various measures of real-world physical behavior were not significantly correlated with FACT-G total well-being (\(P_{s}\geq.189\); section 5 in Multimedia Appendix 1). Average daily time in stepping bouts ≥1 minute was significantly correlated with FACT-G physical well-being (\(p=0.22; \ P=.046\), FACT-G physical well-being 5-item subset (\(p=0.29; \ P=.007\)), and linked PROMIS-PF T scores (\(p=0.29; \ P=.007\), but no other measures of physical behavior were associated with FACT-G physical well-being, FACT-G physical well-being 5-item subset, or linked PROMIS-PF T scores (\(P_{s}\geq.08\); section 5 in Multimedia Appendix 1). When accounting for the effects of demographics and cancer characteristics on these associations using a partial Spearman correlation framework, the pattern of significance was largely unchanged, except that the correlation between time in stepping bouts ≥1 minute and FACT-G physical well-being was no longer significant (Figure S5 in Multimedia Appendix 1).

Individuals with high, medium, and low FACT-G total well-being scores did not differ significantly in terms of any of the measures of real-world physical behavior (Figure 2). Similarly, individuals with high, medium, and low levels of FACT-G physical well-being scores, FACT-G physical well-being 5-item subset scores, and linked PROMIS-PF T scores did not differ significantly in terms of sedentary time, step counts, time in moderate to vigorous activity, weighted median cadence, or peak 30-second cadence (Figure 2; Figure S6 in Multimedia Appendix 1). However, we did find that participants with high FACT-G physical well-being 5-item subset and linked PROMIS-PF T scores spent more time in stepping bouts ≥1 minute than those with medium (\(W=121; \ P=.001\)) and low scores (\(W=155; \ P=.004\)).
Figure 1. Correlation matrix depicting pairwise Spearman correlations with measures of real-world physical behavior. FACT-G: Functional Assessment of Cancer Therapy-General; PROMIS-PF: Patient-Reported Outcomes Measurement Information System-Physical Function. *P<.05, **P<.01, ***P<.001.
Figure 2. Box plots depicting the measures of real-world physical behavior according to tertile splits of self-reported well-being and physical function and a median split of aerobic fitness. FACT-G: Functional Assessment of Cancer Therapy-General; PROMIS-PF: Patient-Reported Outcomes Measurement Information System-Physical Function; submaximal VO₂: submaximal oxygen uptake. Significance labels refer to the results of Welch $t$ tests and Mann-Whitney $U$ tests. For ease of visualization, time in moderate to vigorous activity was transformed with a reverse inverse normal (RIN) transformation.

$^* P < .05$, $^{**} P < .01$, $^{***} P < .001$.

Real-World Physical Behavior Was Associated With Aerobic Fitness

All but one of the accelerometry-derived measures of real-world physical behavior were significantly correlated with submaximal VO₂ ($P$s $\leq .03$; Figure 1; section 5 in Multimedia Appendix 1). Weighted median cadence in stepping bouts $\geq$ 1 minute was the only measure not associated with submaximal VO₂ ($P$=0.08; $P$=.61). After accounting for the effects of demographics and cancer characteristics in a partial Spearman correlation framework (Figure S5 in Multimedia Appendix 1), average...
daily step count was significantly correlated with submaximal VO₂ (p=0.46; P=.002), as was time in moderate to vigorous activity (p=0.33; P=.03).

A median split of submaximal VO₂ (Figure 2) indicated that compared with participants with low submaximal VO₂, participants with high submaximal VO₂ had significantly higher step counts (W=130; P<.001) and spent significantly more time in light intensity activity (t_{42.6}=2.23; P=.03), moderate intensity activity (W=128; P<.001), and stepping bouts ≥1 minute in duration (W=144; P=.002). Individuals with high and low submaximal VO₂ did not differ significantly in terms of sedentary time (W=367; P=.19), weighted median cadence (t_{42.7}=-0.07; P=.95), or peak 30-second cadence (W=212; P=.08).

**Aerobic Fitness Was Not Associated With Self-Reported Well-Being or Physical Function**

Spearman correlation analyses indicated that submaximal VO₂ was not significantly correlated with any of the self-reported measures (Ps≥.21; section 7 in Multimedia Appendix 1). The pattern of significance was unchanged when using a partial correlation approach to account for the effects of demographic and cancer characteristics on these associations (Ps≥.27; section 7 in Multimedia Appendix 1).

**Associations With Real-World Physical Behavior Were Stronger for Aerobic Fitness Than for Self-Reported Well-Being or Physical Function**

Having found that the measures of real-world physical behavior were largely uncorrelated with self-reported well-being and physical function but correlated with aerobic fitness, we used likelihood ratio tests to compare these sets of associations (Figure 3; Figure S8 in Multimedia Appendix 1). These analyses indicated that step count was more strongly associated with submaximal VO₂ than with FACT-G total well-being (F₁=12.29; P=.001), FACT-G physical well-being (F₁=18.27; P<.001), FACT-G physical well-being 5-item subset (F₁=16.32; P<.001), and linked PROMIS-PF T scores (F₁=15.72; P<.001). Similarly, time in moderate to vigorous activity was more strongly associated with submaximal VO₂ than with FACT-G total well-being scores (F₁=7.05; P=.01), FACT-G physical well-being (F₁=8.78; P=.005), FACT-G physical well-being 5-item subset (F₁=8.13; P=.007), and linked PROMIS-PF T scores (F₁=9.30; P=.004). Time in stepping bouts ≥1 minute was also more strongly associated with submaximal VO₂ than with any of the self-reported measures (FACT-G total: F₁=4.87; P=.03; FACT-G physical: F₁=8.34; P=.006; FACT-G physical 5-item subset: F₁=7.16; P=.01; linked PROMIS-PF: F₁=5.48; P=.03).

Sedentary time was more negatively associated with submaximal VO₂ than with FACT-G physical well-being (F₁=7.49; P=.009), FACT-G physical well-being 5-item subset (F₁=5.36; P=.03), and linked PROMIS-PF T scores (F₁=7.02; P=.01), but not with FACT-G total well-being scores (F₁=1.93; P=.17). Similarly, time in light activity was more positively associated with submaximal VO₂ than with FACT-G physical well-being (F₁=4.86; P=.03) and linked PROMIS-PF T scores (F₁=5.01; P=.03), but not with FACT-G total well-being scores (F₁=1.57; P=.22) or physical well-being 5-item subset scores (F₁=4.01; P=.05). For weighted median cadence and peak 30-second cadence, associations with submaximal VO₂ were not significantly different than those with any of the participant-reported measures (Ps≥.08; section 8 in Multimedia Appendix 1).

A similar pattern of results was observed when examining relationships with measures of activity fragmentation (Multimedia Appendix 1). Specifically, measures indicating a more fragmented activity pattern were correlated with lower submaximal VO₂ but were largely unrelated to measures of self-reported well-being and physical function (Figure S9 in Multimedia Appendix 1); furthermore, multiple measures of activity fragmentation were significantly more associated with aerobic fitness than with the self-reported measures (Figure S10 in Multimedia Appendix 1).
Discussion

Principal Findings

Amid a digital revolution in medicine, the use of digital health technologies as evidence generation tools in oncology clinical trials and routine cancer care is gaining traction [27,67]. Wearable sensors are increasingly being used for assessing the efficacy of anticancer therapies and for posttreatment monitoring [29], but the clinical utility of measures of real-world behavior derived from these devices remains to be fully characterized. In this study, we examined how measures of real-world physical behavior, captured in real-world environments of cancer...
survivors over a 1-week monitoring period using accelerometry, were related to self-reported and performance outcomes. We found that the volume and patterning of real-world physical behavior were more related to aerobic fitness than to self-reported well-being and physical function.

Previous studies assessing relationships between real-world measures of physical behavior and self-reported well-being and physical function in cancer survivors have reported mixed findings. In a study of prostate cancer survivors, accelerometer-assessed time spent sedentary, time in light activity, and time in moderate to vigorous activity were all associated with global well-being, but only at specific percentiles of well-being [68]. In colon cancer survivors, time spent sedentary was associated with quality of life [69], and among colorectal cancer survivors, time in moderate to vigorous activity was associated with quality of life and physical function [70]. However, one of these studies failed to find a significant association between sedentary time and either quality of life or physical function [70], and in a separate study, neither time in sedentary behavior nor time in moderate to vigorous activity was significantly associated with quality of life in prostate cancer survivors [71]. Our findings are in line with these prior reports of limited relationships with measures of real-world physical behavior and suggest that these measures are more reflective of objective physical capacity than of self-reported well-being and physical function in cancer survivors.

There are several potential explanations for why we did not observe many significant relationships between real-world physical behavior and the self-reported measures. One reason may be that ceiling effects in the self-reported measures limited our ability to detect associations with physical behavior. We observed ceiling effects for FACT-G physical well-being and linked PROMIS-PF T scores, which may be due to selection bias, as well as some participants being far out from diagnosis and treatment at the time of assessment. All participants in studies 1 and 2 had completed treatment, and participants in both the studies had been diagnosed an average of 32 months before data collection. Ceiling effects are a limitation of some participant-reported assessments, including FACT-G, its subscales, and PROMIS-PF short forms [24,54,72,73], with these effects challenging the ability of an assessment to detect changes over time [56]. These effects may be especially relevant when respondents have higher levels of functioning [24,54], which could occur when assessing cancer survivors (1) years out from diagnosis, (2) with cancer types that tend to be detected early, or (3) who experience relatively smaller declines in functioning. As fewer of the real-world physical behavior measures were highly skewed, these measures have the potential to capture aspects of functioning beyond those captured with self-reported measures.

Another reason may be that the real-world measures studied here do not capture the aspects of real-world physical behavior that are most associated with self-reported well-being and physical function. We included a range of measures of real-world physical behavior, with the aim of gaining insights into their differential clinical utility. Step count, time in moderate to vigorous activity, time spent sedentary, time in light activity, and time in stepping bouts ≥1 minute, all demonstrated stronger associations with aerobic fitness than with the self-reported measures, suggesting that these particular measures may offer more insights into individuals’ physical capacity than their well-being and perceived physical function. We found that weighted median cadence and peak 30-second cadence were largely unrelated to aerobic fitness, and their associations with aerobic fitness did not differ from those for the participant-reported measures. It is worth noting that we calculated time and weighted median cadence in stepping bouts ≥1 minute in duration (rather than longer-duration stepping bouts), due to many participants not spending time in longer-duration stepping bouts. As most stepping bouts taken in day-to-day behavior tend to be <1 minute in duration [74], time and cadence in longer-duration stepping bouts may be more informative, but studies of larger samples are needed to examine the clinical utility of these measures. At the same time, with participants spending the most time in short-duration stepping bouts, aspects of gait such as gait speed and variability not explored here may be clinically relevant measures of day-to-day functioning and worth further investigation.

Beyond measures reflecting the absolute volume of physical behavior, we found that measures reflecting a more fragmented pattern of daily activity and sedentary time were negatively correlated with aerobic fitness but were mostly unrelated to self-reported well-being and physical function. Fragmented daily physical activity has been associated with poorer physical function as measured in the clinic, as well as higher fatigueability [63,64]. Additional research is needed to understand whether measures reflecting the fragmentation of real-world physical behavior can provide additional insights into real-world physical function, beyond measures reflecting the absolute volume of physical behavior, in cancer survivors. Taken together, it may be that further research is needed to define and validate measurable concepts and features of real-world physical behavior that are more closely related to perceived physical function.

We note several other important limitations. First, this was a cross-sectional analysis; results may differ if examining relationships with change in real-world physical behavior. Testing whether real-world physical behavior is associated with established measures of physical function and well-being over time will be necessary for establishing clinical validity of these measures. In addition, the sample size was small, with only 49 individuals included in the analyses involving aerobic fitness due to some in-person assessments being suspended during the COVID-19 pandemic. Furthermore, participants were mostly White and female, with high levels of educational attainment, limiting the ecological validity of results. In addition, most participants were diagnosed at cancer stage II or lower, with breast, colon, or rectal cancer, so results may not generalize to survivors of more advanced cancers or of other cancer types. Similarly, this analysis was focused on individuals who had completed treatment, which allowed us to consider questions of clinical utility without the confounding effects of disease and treatment on functioning; however, results may not generalize to individuals undergoing active cancer treatment. Further investigation of real-world physical behavior in larger, more...
representative samples of individuals during and after cancer treatment is warranted.

Another important limitation is that this was a secondary analysis of previously collected data, and so the studies were not designed to test the questions posed in this investigation. Related to this, it is possible that using other self-report and aerobic fitness measures might have yielded different results. Additional work to probe these relationships with other measures may help inform the clinical utility of wearable-derived digital measures in cancer survivorship.

Beyond these limitations, our findings speak to the potential utility of digital measures of real-world physical behavior to contribute to the assessment of functioning in cancer survivorship. That the digital measures did not exhibit many significant relationships with self-reported well-being and physical function suggests that these sets of measures provide different information. Furthermore, real-world physical behavior was significantly associated with submaximal VO\(_2\), if further investigation reveals significant overlap in the clinical utility of these measures, wearable sensors could provide a lower-burden means of capturing information on aerobic fitness. Finally, compared with using any single type of measure, combining participant-reported, performance, and objective real-world measures could provide a more holistic picture of functioning in cancer survivorship [75]. Taking a comprehensive approach to assessing functioning could furthermore increase sensitivity to detect clinical change over time, enabling more efficient discovery of novel anticancer therapeutics or efficacious interventions for cancer survivors. This approach furthermore offers the possibility to better predict clinical outcomes, which could enable earlier disease detection and the personalization of both treatment and survivorship care [76,77]. As they can be captured remotely and passively, digital measures of real-world physical behavior can also enable decentralization of clinical trials, lower patient burden for participation, and facilitate the recruitment of underrepresented populations [78].

Although our findings indicate that digital measures of real-world physical behavior may add value for the measurement of functioning in cancer survivorship, further research is needed to evaluate the relative value and unique contributions of real-world physical behavior and self-reported physical function to the well-being of cancer survivors. Our approach and previous studies have been limited to cross-sectional analyses, but further work assessing how measures of real-world physical behavior relate to established clinical outcomes over time will be important for advancing the appropriate use of digital measures in oncology clinical research [79,80]. There are additional challenges with implementing wearable sensors in clinical populations, including acceptability and feasibility of these devices among participants. In addition, there is a growing regulatory emphasis on patient centricity in the development of clinical outcome assessments, such that digital measures derived from wearable sensors should reflect aspects of health that are meaningful to individuals in the target clinical population of interest [81]. Our findings suggest that digital measures may provide additional insights into physical function beyond those obtained with self-reported assessments, but whether these insights reflect aspects of everyday functioning that are meaningful remains to be determined. Gathering the evidence needed to demonstrate that digital measures are validated, meaningful, and feasible to capture will be important for promoting broad acceptance and proper use of digital measures in oncology clinical research [79,81,82].

Conclusions

Digital health technologies such as wearable sensors are increasingly used in oncology clinical research and offer potential for capturing aspects of real-world functioning in cancer survivors. In this secondary analysis, we investigated the clinical utility of accelerometry-derived measures of real-world physical behavior in a sample of individuals who had completed cancer treatment. We found that several measures of real-world physical behavior were more associated with aerobic fitness, assessed with a submaximal exercise test, than they were with self-reported measures of well-being and physical function. Our findings suggest that in cancer survivors who have completed treatment, measures of real-world physical behavior may be able to complement self-reported measures of well-being and physical function.

Acknowledgments

This research was supported by contributions from the Department of Health and Exercise Science at Colorado State University and in part by the American Cancer Society (131629-MRSG-18-021-01-CPPB). The authors would like to thank Mary Hidde for assisting with data collection and management for both studies and participants for their willingness to contribute their time and efforts for the purpose of research.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SLB was responsible for methodology, formal analysis, software, validation, visualization, and writing the original draft. EG contributed to data curation, project administration, and writing—review and editing. SA and DC contributed to writing—review and editing. IC contributed to the methodology and writing—review and editing. KL participated in methodology, data curation,
and writing—review and editing. HL contributed to conceptualization, funding acquisition, supervision, project administration, resources, data curation, methodology, writing—review and editing.

Conflicts of Interest
SLB, SA, KL, and IC are employees of VivoSense, Inc. DC is the president of FACIT.org. IC is on the Editorial Board of Karger Digital Biomarkers and the Scientific Advisory Board for IMI IDEA FAST and has received fees for lectures and consulting on digital health at ETH Zürich and FHNW Muttenz. All the other authors declare no conflicts of interest.

Multimedia Appendix 1
Supplementary methods and results.
[PDF (Adobe PDF File), 1193 KB - cancer_v10i1e53180_app1.pdf ]

References


36. Bachman et alJMIR CANCER


49. Reciprocal leg movements: PAL knowledge base. PAL Technologies Ltd. URL: https://kb.palt.com/articles/reciprocal-leg-movements/ [accessed 2023-03-10]


Abbreviations
- FACT-G: Functional Assessment of Cancer Therapy-General
- PROMIS-PF: Patient-Reported Outcomes Measurement Information System-Physical Function

©Shelby L Bachman, Emma Gomes, Suvekshya Aryal, David Cella, Ieuan Clay, Kate Lyden, Heather J Leach. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.07.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Improving Concordance Between Clinicians With Australian Guidelines for Bowel Cancer Prevention Using a Digital Application: Randomized Controlled Crossover Study

Tsai-Wing Ow¹,², MBBS; Olga Sukocheva¹, BSc, PhD; Peter Bampton¹, MBBS; Guruparan Iyngkaran³, MBBS, PhD; Christopher K Rayner¹,², MBBS, PhD; Edmund Tse¹,², MBBS, PhD

¹Department of Gastroenterology and Hepatology, Royal Adelaide Hospital, Adelaide, Australia
²Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, Australia
³Department of Gastroenterology and Hepatology, Royal Melbourne Hospital, Melbourne, Australia

Corresponding Author:
Tsai-Wing Ow, MBBS
Department of Gastroenterology and Hepatology
Royal Adelaide Hospital
Port Road
Adelaide, 5000
Australia
Phone: 61 70740000
Fax: 61 70746247
Email: tsai-wing.ow@sa.gov.au
Faculty of Health and Medical Sciences
University of Adelaide
North Terrace
Adelaide, 5005
Australia
Phone: 61 83135208
Email: tsai-wing.ow@sa.gov.au

Abstract

Background: Australia’s bowel cancer prevention guidelines, following a recent revision, are among the most complex in the world. Detailed decision tables outline screening or surveillance recommendations for 230 case scenarios alongside cessation recommendations for older patients. While these guidelines can help better allocate limited colonoscopy resources, their increasing complexity may limit their adoption and potential benefits. Therefore, tools to support clinicians in navigating these guidelines could be essential for national bowel cancer prevention efforts. Digital applications (DAs) represent a potentially inexpensive and scalable solution but are yet to be tested for this purpose.

Objective: This study aims to assess whether a DA could increase clinician adherence to Australia’s new colorectal cancer screening and surveillance guidelines and determine whether improved usability correlates with greater conformance to guidelines.

Methods: As part of a randomized controlled crossover study, we created a clinical vignette quiz to evaluate the efficacy of a DA in comparison with the standard resource (SR) for making screening and surveillance decisions. Briefings were provided to study participants, which were tailored to their level of familiarity with the guidelines. We measured the adherence of clinicians according to their number of guideline-concordant responses to the scenarios in the quiz using either the DA or the SR. The maximum score was 18, with higher scores indicating improved adherence. We also tested the DA’s usability using the System Usability Scale.

Results: Of 117 participants, 80 were included in the final analysis. Using the SR, the adherence of participants was rated a median (IQR) score of 10 (7.75-13) out of 18. The participants’ adherence improved by 40% (relative risk 1.4, P<.001) when using the DA, reaching a median (IQR) score of 14 (12-17) out of 18. The DA was rated highly for usability with a median (IQR) score of 90 (72.5-95) and ranked in the 96th percentile of systems. There was a moderate correlation between the usability of the DA and better adherence (r=0.4; P<.001). No differences between the adherence of specialists and nonspecialists were found, either with the SR (10 vs 9; P=.47) or with the DA (13 vs 15; P=.24). There was no significant association between participants
who were less adherent with the DA (n=17) and their age (P=.06), experience with decision support tools (P=.51), or academic involvement with a university (P=.39).

Conclusions: DAs can significantly improve the adoption of complex Australian bowel cancer prevention guidelines. As screening and surveillance guidelines become increasingly complex and personalized, these tools will be crucial to help clinicians accurately determine the most appropriate recommendations for their patients. Additional research to understand why some practitioners perform worse with DAs is required. Further improvements in application usability may optimize guideline concordance further.

(JMIR Cancer 2024;10:e46625) doi:10.2196/46625

KEYWORDS

colorectal cancer; guidelines; colorectal cancer screening; digital application; questionnaire; application; cancer prevention; prevention; cancer; bowel cancer; surveillance; clinical vignette quiz; usability; Australia

Introduction

Australia’s National Health and Medical Research Council (NHMRC) screening and surveillance guidelines for colorectal cancer have become substantially more complex with their latest revision [1,2]. This is due to a shift toward personalized recommendations through detailed risk stratification based on an individual’s history of polyps or a family history of cancer. As a result, the guidelines now describe up to 230 different screening or surveillance scenarios, requiring clinicians to navigate through multiple tables to determine an appropriate recommendation. While implementing these changes can considerably improve resource use, this complexity may be a barrier to adherence, limiting the benefits of the guidelines [3,4]. Consequently, there is a need for tools to support clinicians using these guidelines. However, few of these tools have been adequately evaluated.

Several approaches have previously been considered to assist clinicians in determining appropriate bowel cancer prevention guideline recommendations. In the United States, where the complexity of polyp surveillance guidelines is the most similar to those of Australia, researchers have primarily focused on developing methods to assist clinicians in determining the appropriate advice, with a particular emphasis on automating the extraction of clinical information from electronic records to determine guideline-concordant recommendations [5-7]. In clinical practice, this resulted in a small but significant improvement in the rate of guideline-concordant recommendations (84.6% vs 77.4%) [7]. In Australia, print-based educational interventions for screening and surveillance, targeted at patients and clinicians, respectively, have had a minimal impact on improving guideline adherence [8,9]. By contrast, a nurse-led decision-making model has been the most successful intervention, increasing the rate of guideline-concordant recommendations from 83% to 97% [10]. Although successful, these options are associated with substantial costs for setup and maintenance and are not easily scalable beyond individual health services. Furthermore, how they perform when applied to the recently revised Australian guidelines is unclear.

Smartphone- or web-based digital applications (DAs) can be developed cheaply and are readily scalable. However, there are limited studies evaluating their effectiveness in supporting clinician adherence to complex bowel cancer prevention guidelines. Khan et al [11] showed that a DA was able to improve medical students’ knowledge of US colorectal cancer screening guidelines. However, their study was not randomized and did not control for the improvement in scores merely due to repeated exposure to the same clinical questions. In another study, a DA was evaluated by 6 endoscopists assessing a total of 58 colonoscopies [12]. As this was a small pilot study primarily focused on assessing the attitudes of potential users to guide the development of a new DA, it is difficult to draw meaningful conclusions about the potential benefit of the tool in improving guideline concordance.

In Australia, some of the DAs developed in response to the complexity of the latest surveillance guidelines include polyp.guide, polyp.app, and CRCwebapp [13-15]. These 3 tools provide greater ease of use by not requiring users to work through the risk tables manually. To the best of our knowledge, only CRCwebapp has been validated against all 230 possible case scenarios due to its use as a research tool in a previously published study [3]. However, none of these have been evaluated for their ability to improve the rate of guideline concordance among clinicians.

We hypothesized that a DA could improve clinician adherence to Australian screening and surveillance guidelines. To test this, we conducted a randomized controlled crossover study to compare the proportion of guideline-concordant decisions made by clinicians using either the CRCwebapp DA or the standard resource (SR).

Methods

Study Design and Setting

We enrolled practicing Australian clinicians to our online randomized controlled crossover clinical vignette questionnaire between July 1, 2020, and August 1, 2021. Participants were asked to provide guideline-concordant recommendations for 2 sets of clinical vignettes using either the SR or the DA. All participants were provided with an orientation that was tailored according to their experience with the guidelines. The clinical vignettes and order in which the tools were used were randomized. A study portal was used to present the vignettes, and this provided participants with access to both the SR and DA. After completing questions related to the clinical vignettes with both the SR and DA, the System Usability Scale (SUS) questionnaire was administered.
Inclusion Criteria
We included medical, surgical, or specialist nurse practitioners who were actively practicing in Australia during the study period.

Exclusion Criteria
Participants who were not actively involved in making screening or surveillance decisions for colorectal cancer in their clinical work were excluded.

Participant Orientation
We classified participants into 2 groups according to their familiarity with the guidelines. The nonspecialist group comprised primary care practitioners who had limited experience with the terminology and structure of the published guidelines. The specialist group comprised gastroenterologists, colorectal surgeons, and specialist nurse practitioners who were routinely using the current screening and surveillance guidelines in clinical practice. The orientation program was tailored according to the experience of each group, in order to reduce the impact of experience on participant scores and to reduce barriers to participation.

For nonspecialists, the necessary terminology pertaining to screening and surveillance was defined during a web seminar. This included degree of relationship in family history for screening protocols and the individual risk characteristics and classification of lesions for surveillance protocols. The seminar also included a breakdown of every decision table in the SR and the most efficient methods to navigate to each of these. Participants were also introduced to the 4 main pages of the DA and shown how to input data and where the results were presented. In contrast, the specialist orientation did not define the terminology, and the introductions to the SR and the DA were presented as optional videos available before the questionnaire.

Primary Outcome
The primary outcome was the proportion of correct screening and surveillance recommendations issued by participants in response to the clinical vignettes. Each vignette could receive a maximum score of 6, resulting in each participant being graded with a score out of 18 for each of 2 sets of 3 clinical vignettes.

Secondary Outcome
The secondary outcome was the usability of the DA. This was assessed using each participant’s response to the SUS. A score was determined for each participant and normalized in accordance with previously published methods [16].

Clinical Vignette Design
Three pairs of clinical vignettes were developed for the study (alpha and beta, gamma and theta, and delta and omega). Each vignette described the family history, medical comorbidities, and the number and characteristics of conventional adenomas or sessile serrated lesions identified over the preceding 2 colonoscopies. We avoided scenarios commonly highlighted in previous guidelines to reduce the likelihood that participants could answer according to their recollection of these [17]. Each pair of clinical vignettes focused participants on navigating identical sets of tables to balance for difficulty.

For each clinical vignette, participants were asked to determine the age and appropriate screening modality (stool testing or colonoscopy) based on the family history presented, the first and subsequent recommended surveillance intervals, and whether surveillance should be continued when considering the comorbidities of the patient if the age of the patient was >75 years at the time of the intended procedure. Each vignette received a score out of 6. Thus, each participant could receive a maximum score of 18 for each section.

Usability
We adapted the standard SUS questionnaire by changing the term “system” to “application” in order to focus participants on assessing the usability of the DA (Textbox 1). This comprised 10 standardized statements for which users were asked to indicate their level of agreement. Numerical scores provided by participants on a slider scale were translated into Likert scores: 0-20=strongly disagree (1); 21-40=disagree (2); 41-60=neither agree nor disagree (3); 61-80=agree (4); and 81-100=strongly agree (5). A total SUS score was calculated for each participant [18]. The scores were normalized to provide a percentile ranking of the usability of the DA, as described by Sauro and Lewis [16].

Textbox 1. System Usability Scale questionnaire adapted for the use of the digital application.

1. I think that I would like to use this application frequently.
2. I found the application unnecessarily complex.
3. I thought the application was easy to use.
4. I think that I would need the support of a technical person to be able to use this application.
5. I found that the various functions in this application were well integrated.
6. I thought there was too much inconsistency in this application.
7. I would imagine that people would learn to use this application very quickly.
8. I found the application very cumbersome to use.
9. I felt very confident using the application.
10. I needed to learn a lot of things before I could get going with this application.
DA Design

Each NHMRC screening and surveillance recommendation was coded into an Excel (Microsoft Corp) spreadsheet. We eliminated redundant user data entry by determining the minimum number of inputs necessary to calculate each recommendation. For screening decisions, this included 4 fields relating to the number and age of relatives with colorectal cancer and their relation (first or second degree) to the patient. For surveillance intervals, this included the number, type, and characteristics of the lesions found during the initial procedure. Subsequent surveillance intervals required 2 additional inputs: the initial surveillance interval and the type of lesion previously identified. An additional section, incorporating a list of potential patient comorbidities, was used to determine stopping rules.

A graphical user interface was applied using an open-source platform (Open as App), which would allow for the distribution of the DA as either a web page or smartphone app. Each type of calculation (screening, first surveillance, second surveillance, or stopping rules) was identified by a tab on the bottom of the screen. Sliders were used to input data on the number of lesions, and drop-down menus were used to provide details regarding the accompanying risk characteristics. The recommendations for screening, surveillance interval, or cessation of surveillance were provided at the bottom of each respective page. The answers provided by the digital calculator were validated by individually calculating all possible scenarios covered by the updated guidelines before recruitment.

SR for Screening and Surveillance

The SR was the official web publication of the latest guidelines for screening and surveillance for bowel cancer prevention in Australia by the NHMRC [1,2]. In addition to a written summary, it provides details regarding the development of and evidence for each recommendation. Also included are a series of colored risk stratification tables to guide users through screening, initial and follow-up surveillance, and stopping rules. For screening, 90 possible scenarios are defined according to the number of relatives with colorectal cancer as well as how closely they are related to the patient.

For initial surveillance colonoscopy, 37 separate scenarios are described across 3 tables according to the various combinations of “conventional adenomas” or “clinically significant serrated polyps” identified. A total of 140 scenarios are similarly characterized across an additional 9 tables to account for the possible combinations of “conventional adenomas” and “clinically significant serrated polyps” between 2 consecutive procedures. Determining the correct surveillance interval can thus require users to successfully navigate 2 consecutive tables.

Lastly, the rules for cessation of surveillance colonoscopy are detailed in a text table that uses a modified Charlson score. Scores are allocated according to age and the presence of comorbidities. Depending on the combination of age and severity of comorbid conditions, the benefit of continuing surveillance for patients may be deemed too low to justify the potential risks of colonoscopy.

Recruitment

Advertising flyers were created and distributed to the 3 local Primary Health Networks, social media (Facebook: Adelaide GP Referral Network, Medical Mums, and Mums To Be), general practitioner education providers (GPEx and GP Synergy), and directly to practice managers located within metropolitan Adelaide. Additional flyers for specialists were distributed to members of the Departments of Gastroenterology and Hepatology and Colorectal Surgery Departments at 4 major teaching hospitals in Adelaide, as well as to private specialist practices. Snowball sampling was used to aid in the recruitment of additional participants. Continuing professional development points and a certificate of completion were awarded as an incentive to improve recruitment.

Data Collection

The questionnaire was programmed using REDCap (Research Electronic Data Capture; Vanderbilt University) tools hosted at the University of Technology, Sydney, and accessed through the Australian Access Federation [19,20]. We collected data including each participant’s age, professional background (general practice, medical specialist or trainee, and surgical specialist or trainee), active affiliation with a university, and experience with tools supporting screening and surveillance guidelines. We scored the answers for each clinical vignette in the order in which they were completed and collected each participant’s responses to the SUS questionnaire regarding their experience with the DA on a digital spreadsheet for analysis according to a previously described methodology [15].

Randomization

Two randomly permuted schedules (primary care and specialist groups) were created for a crossover study with 2 interventions (DA vs SR) with equal allocation over 8 strata (combinations 1-8; Table 1). A total of 14 allocations were generated per stratum with a total of 112 allocations. Participants were randomized to use either the DA or SR as the first aid in a 1:1 ratio. The 2 allocation schedules were programmed into the REDCap software using branching logic tools. The randomization schema was generated using Microsoft Excel (version 16.66.1; Microsoft Corp).
Table 1. Clinical vignette combinations used for randomization.

<table>
<thead>
<tr>
<th>Combination</th>
<th>Section 1</th>
<th>Section 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alpha, gamma, and delta</td>
<td>Beta, theta, and omega</td>
</tr>
<tr>
<td>2</td>
<td>Alpha, gamma, and omega</td>
<td>Beta, theta, and delta</td>
</tr>
<tr>
<td>3</td>
<td>Alpha, theta, and delta</td>
<td>Beta, gamma, and omega</td>
</tr>
<tr>
<td>4</td>
<td>Alpha, theta, and omega</td>
<td>Beta, gamma, and delta</td>
</tr>
<tr>
<td>5</td>
<td>Beta, gamma, and delta</td>
<td>Alpha, theta, and omega</td>
</tr>
<tr>
<td>6</td>
<td>Beta, gamma, and omega</td>
<td>Alpha, theta, and delta</td>
</tr>
<tr>
<td>7</td>
<td>Beta, theta, and delta</td>
<td>Alpha, gamma, and omega</td>
</tr>
<tr>
<td>8</td>
<td>Beta, theta, and omega</td>
<td>Alpha, gamma, and delta</td>
</tr>
</tbody>
</table>

Statistics

Previously reported rates of adherence to Australian surveillance guidelines have ranged from 50.8% to 83% [4,10]. The impact of a nurse-led intervention improved the rate of guideline concordance by a factor of 1.17 relative to the non–nurse-led group [10]. On the basis of these results, we predicted a mean accuracy score of 60% with the SR and anticipated a 1.17 improvement in the rate of guideline concordance to 70% with the intervention (DA). Using an expected SD of 20%, an \( \alpha \) of .05, and a statistical power of 0.8, the minimum necessary sample size required was calculated at 64 participants.

Descriptive statistics were used to characterize the data. A Kolmogorov-Smirnov test was applied to assess for normality of the data before the statistical analysis. A related-samples Wilcoxon signed rank test was used to compare the performances of participants with either the SR or the DA. An independent-samples Mann-Whitney \( U \) test was used to compare outcomes between specialists and nonspecialists. Spearman \( \rho \) was used to assess the relationship between usability and scores from the DA. \( \chi^2 \) tests of independence were used to compare the allocation of participants between tools and clinical vignettes. The SPSS statistical software (version 22; IBM Corp) was used for all analyses.

Ethical Considerations

The study protocol was reviewed and approved by the Central Adelaide Local Health Network Human Research Ethics Committee (CALHN Research Office reference 13438). The background, procedures, and aims of the study were provided to prospective participants via a digital participant information sheet before the commencement of the survey. Participants were informed that their consent to participate would be implied via completion and submission of the online questionnaire. All data collected were deidentified. No participants received financial compensation.

Results

Participant Characteristics

In total, 117 participants initiated the questionnaire. The records of 37 participants were excluded from the primary analysis due to survey noncompletion. Of these, no components of the questionnaire were attempted in 8 cases. 25 participants completed the background survey but did not attempt the clinical vignette section, and 4 participants aborted the clinical vignette section before completion (Figure 1). These included 7 primary care doctors, 20 gastroenterologists, 1 surgeon, 1 nurse endoscopist, and 8 participants of unknown vocation. One additional participant aborted the study after completing the vignettes and was included in the primary analysis but not in the evaluation of the usability scores.
The remaining 80 participants, consisting of 43 primary care doctors and 37 specialist doctors (35 gastroenterologists and 2 surgeons), were included in the primary analysis. They had a median age of 38 (IQR 27-71) years. Fewer than half (35/80, 44%) held an affiliation with a university (27/37, 73% of specialists and 8/43, 19% of primary care doctors), and almost two-thirds (51/80, 64%) had previously used tools for screening and surveillance decisions in colorectal cancer (32/37, 87% of specialists and 19/43, 44% of primary care doctors; Table 2). The study flowchart shows how participants were randomized to 1 of 8 sequences of vignettes (Figure 1). Of the 80 included participants, 38 (48%) were assigned to use the DA as the first aid (Figure 2). Alpha, gamma, and delta were the first vignettes in their respective pairs in 48% (38/80), 56% (45/80), and 51% (41/80) of cases (Figure 2).
Figure 2. Allocation of tools and vignettes for the first set of clinical vignettes after randomization. Pearson $\chi^2$ tests of independence were used to assess the distribution order of tools (standard resource or digital application) and vignettes (alpha or beta, gamma or theta, and delta or omega) after excluding participants who did not complete the study. The analysis confirmed that the differences in the final allocation of participants at each stage after exclusions were not significant.
Table 2. Tools to aid decisions in colorectal cancer screening and surveillance (N=80).

<table>
<thead>
<tr>
<th>Tool</th>
<th>Specialist (n=37), n (%)</th>
<th>Primary care (n=43), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiki.cancer Guideline (NHMRC&lt;sup&gt;a&lt;/sup&gt;)</td>
<td>20 (54)</td>
<td>11 (26)</td>
</tr>
<tr>
<td>Polyp.guide</td>
<td>8 (22)</td>
<td>3 (7)</td>
</tr>
<tr>
<td>Digital calculator</td>
<td>6 (16)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Media in endoscopy suite</td>
<td>19 (51)</td>
<td>__&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Polyp nurse support</td>
<td>3 (8)</td>
<td>__</td>
</tr>
<tr>
<td>Funding codes (Medicare)</td>
<td>2 (5)</td>
<td>__</td>
</tr>
<tr>
<td>The Royal Australian College of General Practitioners’ Redbook</td>
<td>__&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6 (14)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (3)</td>
<td>2 (5%)</td>
</tr>
</tbody>
</table>

<sup>a</sup>NHMRC: National Health and Medical Research Council.

<sup>b</sup>Not available.

A Kolmogorov-Smirnov test of normality indicated that the scores of participants using the SR were normally distributed: D(80)=0.075; P=.20, while those of the DA were not: D(80)=0.152; P<.001. With the SR, the median (IQR) number of guideline concordant answers was 10 (7.75-13) out of 18. The use of the DA improved the number of correct recommendations to a median (IQR) of 14 (12-17) out of 18 (relative risk 1.4, P<.001; Figures 3 and 4). Lower performance with the DA compared with SR (n=17) was not associated with previous experience with screening and surveillance decision tools (P=.51), affiliation with a university (P=.39), or age (P=.06).

**Figure 3.** Comparison of spread of clinical vignette scores with either the standard resource (SR) or the digital application (DA). The participant scores when using the SR showed a normal distribution. A rightward shift in the distribution of the scores was observed with the use of the DA.
Figure 4. Box and whisker plot of clinical vignette scores with either the standard resource or the digital application.

The median (IQR) SUS score for the DA was 90 (72.5-95), which equated to a top 4 percentile ranking among tested applications (Table 3). A moderate correlation between usability grade and DA results was observed using Spearman ρ correlation coefficient ($r_s=0.4; P<.001; n=79$).

Table 3. System Usability Scale (SUS) grades and percentiles for participants using the digital application (n=79).

<table>
<thead>
<tr>
<th>Grade</th>
<th>SUS</th>
<th>Participants, n (%)</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>&gt;78.8</td>
<td>51 (65)</td>
<td>85-100</td>
</tr>
<tr>
<td>B</td>
<td>72.6-78.8</td>
<td>7 (9)</td>
<td>65-84</td>
</tr>
<tr>
<td>C</td>
<td>62.7-72.5</td>
<td>11 (14)</td>
<td>35-64</td>
</tr>
<tr>
<td>D</td>
<td>51.7-62.6</td>
<td>6 (8)</td>
<td>15-34</td>
</tr>
<tr>
<td>F</td>
<td>0-51.7</td>
<td>4 (5)</td>
<td>0-15</td>
</tr>
</tbody>
</table>

Sensitivity Analysis

After excluding those who did not complete the study, differences in the randomization of participants regarding order of use of the tools (SR vs DA) and clinical vignettes (alpha vs beta, gamma vs theta, and delta vs gamma) were not significant (Figure 2). Additionally, there was no difference ($P=.55$) between the median number of guideline concordant recommendations according to whether the clinical vignettes were posed to participants: first (12, IQR 8.75-15) or second (13, IQR 9-16; Figure 5). Similarly, no difference was observed between the performance of specialists and primary care doctors, either with the SR (10 vs 9; $P=.47$) or with the DA (13 vs 15, $P=.24$; Figures 6 and 7).
Figure 5. Box and whisker plot of clinical vignette scores according to the order they were answered (first or second). A related-samples Wilcoxon signed rank test was used to compare the results achieved in the first and second set of questions indicating no significant difference ($P=.55$). Thus, increasing familiarity with the format of the questionnaire did not improve the scores achieved by participants.

Figure 6. Box and whisker plot of clinical vignette scores according to the vocational training of the participants (specialist or primary care) using the standard resource (SR). An independent-samples Mann-Whitney U test was used to compare the results of specialists with primary care doctors using the SR. There was no significant difference ($P=.47$) in the performance of participants based on their previous training in either specialist or primary care.
Figure 7. Box and whisker plot of clinical vignette scores according to the vocational training of the participants (specialist or primary care) using the digital application (DA). An independent-samples Mann-Whitney U test was used to compare the results of specialists with primary care doctors using the DA. There was no significant difference ($P=.24$) in the performance of participants based on their previous training in either specialist or primary care.

Discussion

Principal Findings

The findings of this study showed that the adherence of clinicians with Australia’s current screening and surveillance guidelines in their current form is limited. This was significantly improved when clinicians used a DA to assist their navigation of these complex guidelines. These findings were independent of the clinicians’ level of specialization, age, university affiliation, or experience with the use of other decision support tools. However, greater adherence was associated with better DA usability ratings, highlighting the importance of this attribute as a potential target to further bolster clinician guideline adherence.

Australia’s screening and surveillance guidelines are among the most complex worldwide. With the increasing trend toward personalized health care and our growing knowledge of colorectal cancer risk factors, guidelines are likely to continue increasing in complexity. For clinicians, navigating these guidelines in busy practices can be challenging. Even under the controlled conditions of our testing environment, participants could only provide appropriate recommendations in slightly over half of the questions when evaluating the scenarios only with the SR. These findings are consistent with another recent report that assessed the concordance of surveillance recommendations with current guidelines [4]. Because the adherence to previous relatively more straightforward guidelines was already known to be suboptimal, it could be anticipated that rates of adherence may be even lower as their complexity increases. This could undermine their potential benefits in the care of patients and the allocation of limited colonoscopy resources in Australia.

DAs can play an important role in supporting the implementation of Australia’s complex bowel cancer prevention guidelines. Not only do they improve the ability of clinicians to provide guideline-concordant recommendations, as demonstrated by our study, but they can be developed at a relatively low cost and are scalable to a national level. Furthermore, they can be updated with future revisions of the guidelines, ensuring that clinicians can continue to make decisions that are in keeping with the latest evidence.

Despite their clear advantages, the role of DAs in supporting complex guideline adoption has received little attention in the literature. To date, only 2 studies have evaluated DAs in assisting medical personnel with the application of bowel cancer screening and surveillance guidelines. However, these were assessed in relation to US guidelines and are limited by their small size and lack of a randomized controlled methodology. To our knowledge, our study is the first to evaluate a DA using a rigorous randomized controlled crossover design.

Participants provided discordant recommendations in 22% of clinical decisions despite assistance from the DA. However, as the DA used in this study had been validated across all the possible scenarios provided by the guidelines, we considered other factors that may have contributed to this. Our results showed that poor performance with the DA relative to the SR was not associated with participant age, academic experience, or prior experience with similar tools. One area that may have
contributed was DA usability. Although the DA scored very well in the SUS, ranking at or above the 96th percentile of tested systems, there was still a relatively large spread of scores (median 90, IQR 72.5-95) and a moderate correlation between SUS scores and participant performance. This suggests that improvements directed at improving usability for those who scored the DA less well could bolster the adherence rate of clinicians with guidelines; however, the magnitude of overall improvement may be small. Therefore, additional research to gather the opinions of participants who found the interface difficult to use and quantify the degree of progress achieved by addressing these is required.

Human error is another potential factor contributing to the rate of discordant answers. Despite simplifying the process of determining guideline-concordant recommendations, the DA still requires individuals to extract relevant and appropriate data from sometimes complex patient histories. Although human error remains an inevitable component of any interface requiring human input, natural language processing software, which has been used in prior US-based studies, could provide a valuable adjunct to a mobile app [5,6]. This would retain the scalability and portability of the DA but would require additional research, development, and testing before it could be implemented. Such a tool could provide a better balance of the advantages of the tools tested thus far.

**Strengths**

Our study design accounted for the possibility that participants could improve their performance in the clinical vignettes simply due to increasing experience with the questionnaire, by randomizing the order of use of the 2 aids (SR or DA). Furthermore, although the clinical vignettes were designed in pairs that were balanced for difficulty, the order in which each pair was presented to the participant was randomized to limit the risk of bias. The vignettes also focused on clinical scenarios with updated and distinct recommendations within the guidelines, requiring participants to determine the correct answer solely through navigation of the SR or DA.

Another strength of our study was the ability to cater for participants with varying levels of familiarity with guidelines. Our participants included specialists, who are accustomed to using colorectal cancer screening and surveillance guidelines in their everyday practice, and nonspecialists, whose breadth of clinical practice typically limits their experience with specialty guidelines. As these differences may have impacted participant vignette questionnaire scores, particularly those encountered without the DA, we tailored the introductory briefings to provide nonspecialists with additional information in the structured seminars. When the 2 groups were compared, no significant differences between them were observed, either with the SR or with the DA. Although this indicated that the potential effect of experience had been controlled for during our study, it is not possible to say whether this resulted from our differential approach to participant briefing, as this was not an outcome that was measured during our study.

Additionally, we were able to control successfully for potential confounders by randomizing both the order of the questions and the tools used by participants during the vignettes. This was used to address the possibility that participant scores may have improved over time and that the clinical vignettes may not have been completely balanced in their difficulty. Our sensitivity analysis showed that the distribution of questions and tools remained balanced, even after exclusion of participants, and that increasing participant experience with the questionnaire did not result in higher scores.

**Limitations**

While the vignettes intentionally challenged participants to navigate the breadth of the decision tables, only a limited number of scenarios are typically encountered in clinical practice. More than 95% of patients will be classified in the lowest risk category for screening based on family history, while most colonoscopies in Australia will detect few or no significant lesions [1]. Thus, participant performance in this study may not be indicative of real-world application. However, adherence rates to current surveillance guidelines, which have been reported at 50.8%, closely resemble the scores obtained using the SR in our study [4]. Clarification of the real-world efficacy of the DA will require further studies, for example, through prospective randomized nested case-controlled studies involving both primary and specialist group practices.

Our study was also prone to sampling bias. Despite efforts to circulate advertising material for the study via social media, education providers, and hospitals, only 117 participants visited the questionnaire website, and the recruitment rate was slow. The diversity of our sample group was also affected, with surgeons outnumbered by gastroenterologists in the specialist group (2-35). Due to our specific subject matter, it is possible that our participants held favorable views toward technology that may not be representative of the greater community of medical professionals. Although these challenges are not uncommon among studies recruiting clinical personnel as participants, the generalizability of our findings may be limited [21].

Finally, as both resources were readily available for public access at the time of the study, it was not possible to restrict participants to using the tools in the prerandomized order specified. Our intention-to-treat analysis may therefore have underestimated the potential differences in the scores obtained by users in the trial.

**Conclusions**

Australia’s bowel cancer screening and surveillance guidelines have become increasingly complex, posing a challenge for clinicians trying to make appropriate recommendations. Currently, the available options to assist them are costly and need more scalability. DAs represent an inexpensive and scalable solution that enhances guideline concordance among clinicians. Further development and assessment of these tools could improve screening and surveillance outcomes and optimize resource use in an era of increasingly complex and personalized care.
Acknowledgments
We would like to acknowledge the participants who kindly donated their time for this study, without which it would not have been possible.

Data Availability
The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors’ Contributions
TO was responsible for study conceptualization, data curation, formal analysis, investigation, methodology, project administration, programming of the online questionnaire and digital application, and writing and review of the manuscript. OS was involved in the formal analysis and review and editing of the manuscript. PB and GI were involved in the conceptualization of the study. CR and ET were involved in the conceptualization, supervision, and review and editing of the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
CONSORT checklist.
[PDF File (Adobe PDF File), 70 KB - cancer_v10i1e46625_app1.pdf]

References

Abbreviations

DA: digital application
NHMRC: National Health and Medical Research Council
REDCap: Research Electronic Data Capture
SR: standard resource
SUS: System Usability Scale

©Tsai-Wing Ow, Olga Sukocheva, Peter Bampton, Guruparan Iyngkaran, Christopher K Rayner, Edmund Tse. Originally published in JMIR Cancer (https://cancer.jmir.org), 22.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Implementation of Health IT for Cancer Screening in US Primary Care: Scoping Review

Constance Owens-Jasey\textsuperscript{1,2,3}, MPH, PhD; Jinying Chen\textsuperscript{4,5,6}, PhD; Ran Xu\textsuperscript{7}, MD; Heather Angier\textsuperscript{8}, MPH, PhD; Amy G Huebschmann\textsuperscript{9}, MSc, MD; Mayuko Ito Fukunaga\textsuperscript{10}, MD; Krisda H Chaiyachati\textsuperscript{11,12,13}, MPH, MSPH, MD; Katharine A Rendle\textsuperscript{11}, MSW, MPH, PhD; Kim Robien\textsuperscript{14}, PhD; Lisa DiMartino\textsuperscript{15,16}, MPH, PhD; Daniel J Amante\textsuperscript{17}, MPH, PhD; Jamie M Faro\textsuperscript{17}, PhD; Maura M Kepper\textsuperscript{18}, PhD; Alex T Ramsey\textsuperscript{19}, PhD; Eric Bressman\textsuperscript{13,20}, MSHP, MD; Rachel Gold\textsuperscript{1,3,21}, MPH, PhD

\textsuperscript{1}BRIDGE-C2 Implementation Science Center in Cancer Control, Oregon Health & Science University, Portland, OR, United States
\textsuperscript{2}Department of Health Administration and Policy, College of Public Health, George Mason University, Fairfax, VA, United States
\textsuperscript{3}OCHIN, Inc, Portland, OR, United States
\textsuperscript{4}Department of Preventive Medicine and Epidemiology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States
\textsuperscript{5}Data Science Core, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States
\textsuperscript{6}iDAPT Implementation Science Center for Cancer Control, Wake Forest School of Medicine, Winston Salem, NC, United States
\textsuperscript{7}Department of Medicine, Massachusetts General Hospital, Boston, MA, United States
\textsuperscript{8}Department of Family Medicine, Oregon Health & Science University, Portland, OR, United States
\textsuperscript{9}Adult and Child Center for Outcomes Research and Delivery Science, Ludeman Family Center for Women’s Health Research, Division of General Internal Medicine, University of Colorado School of Medicine, Aurora, CO, United States
\textsuperscript{10}Department of Medicine, UMass Chan Medical School, Worcester, MA, United States
\textsuperscript{11}Verily Life Sciences, South San Francisco, CA, United States
\textsuperscript{12}Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, United States
\textsuperscript{13}Milken Institute School of Public Health, George Washington University, Washington, DC, United States
\textsuperscript{14}RTI International, Research Triangle Park, NC, United States
\textsuperscript{15}UT Southwestern Medical Center, University of Texas, Dallas, TX, United States
\textsuperscript{16}Department of Population and Quantitative Health Sciences, UMass Chan Medical School, Worcester, MA, United States
\textsuperscript{17}Brown School, Washington University, St. Louis, MO, United States
\textsuperscript{18}Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, United States
\textsuperscript{19}Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
\textsuperscript{20}Kaiser Permanente Center for Health Research, Portland, OR, United States

Corresponding Author:
Constance Owens-Jasey, MPH, PhD
OCHIN, Inc
PO Box 5426
Portland, OR, 97228-5426
United States
Phone: 1 503 943 2500
Email: owensc@ochin.org

Abstract

Background: A substantial percentage of the US population is not up to date on guideline-recommended cancer screenings. Identifying interventions that effectively improve screening rates would enhance the delivery of such screening. Interventions involving health IT (HIT) show promise, but much remains unknown about how HIT is optimized to support cancer screening in primary care.

Objective: This scoping review aims to identify (1) HIT-based interventions that effectively support guideline concordance in breast, cervical, and colorectal cancer screening provision and follow-up in the primary care setting and (2) barriers or facilitators to the implementation of effective HIT in this setting.
Methods: Following scoping review guidelines, we searched MEDLINE, CINAHL Plus, Web of Science, and IEEE Xplore databases for US-based studies from 2015 to 2021 that featured HIT targeting breast, colorectal, and cervical cancer screening in primary care. Studies were dual screened using a review criteria checklist. Data extraction was guided by the following implementation science frameworks: the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework; the Expert Recommendations for Implementing Change taxonomy; and implementation strategy reporting domains. It was also guided by the Integrated Technology Implementation Model that incorporates theories of both implementation science and technology adoption. Reporting was guided by PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews).

Results: A total of 101 studies met the inclusion criteria. Most studies (85/101, 84.2%) involved electronic health record–based HIT interventions. The most common HIT function was clinical decision support, primarily used for panel management or at the point of care. Most studies related to HIT targeting colorectal cancer screening (83/101, 82.2%), followed by studies related to breast cancer screening (28/101, 27.7%), and cervical cancer screening (19/101, 18.8%). Improvements in cancer screening were associated with HIT-based interventions in most studies (36/54, 67% of colorectal cancer–relevant studies; 9/14, 64% of breast cancer–relevant studies; and 7/10, 70% of cervical cancer–relevant studies). Most studies (79/101, 78.2%) reported on the reach of certain interventions, while 17.8% (18/101) of the included studies reported on the adoption or maintenance. Reported barriers and facilitators to HIT adoption primarily related to inner context factors of primary care settings (eg, staffing and organizational policies that support or hinder HIT adoption). Implementation strategies for HIT adoption were reported in 23.8% (24/101) of the included studies.

Conclusions: There are substantial evidence gaps regarding the effectiveness of HIT-based interventions, especially those targeting guideline-concordant breast and colorectal cancer screening in primary care. Even less is known about how to enhance the adoption of technologies that have been proven effective in supporting breast, colorectal, or cervical cancer screening. Research is needed to ensure that the potential benefits of effective HIT-based interventions equitably reach diverse primary care populations.

(KEYWORDS cancer prevention; health information technology; implementation; implementation strategies; scoping review)

Introduction

Background

For common cancer types such as cervical, colorectal, and breast cancer, routine screening provided in primary care settings can save lives [1]. Although evidence-based national guidelines exist for the provision of such screenings [1-4], patient receipt of guideline-concordant cancer screening is suboptimal nationally and varies substantially across clinical settings [5,6]. This is driven by multiple factors, including provider-level barriers such as the challenge of staying current on changing cancer screening guidelines [6] and the cognitive overload that providers can face when managing the needs of patients with complex conditions [7-11]. Patient-level barriers include lack of knowledge of screening recommendations [6], loss to follow-up [12], fear about screening procedures or outcomes, and financial and logistical challenges [13].

Understanding which interventions effectively address these challenges—and the barriers and facilitators to implementing such interventions—is needed to enhance the delivery of guideline-concordant cancer screening in primary care. The Community Preventive Services Task Force summary of evidence-based interventions for addressing barriers to guideline-concordant cancer screening [14] identifies health IT (HIT)–based interventions as showing particular promise [15-17]. Prior systematic reviews found that HIT-based interventions such as patient reminders and provider feedback tools can be effective in supporting cancer prevention care [15,17,18]. Such interventions can enhance provider-patient communication about cancer screening [19-22]. These interventions can also help care teams identify patients due for screening with automated reminders embedded in the electronic health record (EHR) that can appear either at the point of care [23] and during panel or population management [24].

Yet HIT-based interventions targeting numerous health outcomes are underused in primary care settings [23,25]. One recent systematic review involving 55 studies showed that clinical decision support tools were adopted in <35% of eligible encounters [26]. The adoption of such interventions is impeded by multilevel barriers, such as the challenges inherent to integrating new tools into clinical workflows [27], and lack of training in how to use such tools [28,29]. There is a need to understand best practices for enhancing the adoption of effective HIT-based interventions targeting cancer prevention, including how barriers to the adoption of such interventions can best be addressed in primary care [17,18,30,31].

Objectives

In 2020, the National Cancer Institute’s Consortium for Cancer Implementation Science (CCIS) “Technology in Implementation Science Action Group” identified a need for the scoping review presented here. This review aims to describe the specific knowledge gaps in this evidence base, that is, what is known and unknown about the implementation of effective HIT for cancer screening in primary care. Specifically, it aims to identify (1) HIT-based interventions that effectively support guideline concordance in breast, cervical, and colorectal cancer screening provision and follow-up in the primary care setting and (2) barriers or facilitators to the implementation of effective HIT in this setting. To refine the scope of this review, we focused...
on common cancer screenings that are in the purview of primary care: breast, colorectal, and cervical cancer screening. We note that earlier systematic reviews [15,17,18] assessed the effectiveness of HIT-based interventions at improving cancer screening rates in primary care, but the most recent included data up to June 2014 [15]. This review first summarizes related evidence accrued since 2014 and then assesses current knowledge on the adoption of such interventions. To our knowledge, this is the first scoping review to assess the implementation of HIT in cancer screening.

Methods

Overview
This scoping review was conducted by a multidisciplinary team of researchers from the CCIS with expertise in implementation science, health informatics, health services research, and cancer control. We followed the 6-stage scoping review methodology described by Arksey and O’Malley [32], with consideration of later modifications to this approach made by Levac et al [33]. This review was reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [34].

Ethical Considerations
Ethics approval from the George Mason University Institutional Review Board was not required for this review.

Research Questions
This scoping review was designed to answer two overarching questions: (1) What is known about how HIT-based interventions are used to enhance guideline concordance of cancer screening in primary care settings? (2) What is known about the barriers or facilitators to the implementation and dissemination of these interventions?

Identifying Relevant Studies
With assistance from a health sciences librarian, the first author (COJ) conducted a 3-step search process to identify relevant US-based peer-reviewed and gray literature studies. First, the following bibliographic databases were systematically searched: MEDLINE, CINAHL Plus, Web of Science, and IEEE Xplore. These databases were searched using a combination of search strings that included relevant controlled vocabulary (eg, Medical Subject Heading) and keywords with Boolean operators. The search terms were selected based on a review of the existing literature and refined based on the input of the coauthors. To ensure that the search yielded relevant studies, variations of the search strategy were pilot-tested by 3 authors (COJ, RG, and RX) and refined before the final search was conducted. Our final search strategy for bibliographic databases is provided in Multimedia Appendix 1.

Second, this search was supplemented by a review of gray literature (eg, study protocols, unpublished empirical trials, dissertations, reports, and government publications) to consider studies that might not be indexed in bibliographic databases. This search primarily consisted of targeted website searching of cancer, HIT, public health organizations, and funding agencies recommended by the authors (COJ, RG, KHC). Our final gray literature search strategy is provided in Multimedia Appendix 2. Additional gray literature databases (CQ Press Library, Policy File Index, Find Policy, and Harvard Kennedy School Think Tank Search), recommended by the health sciences librarian, were explored but did not yield useful results. Finally, we identified relevant studies with a snowball search technique, whereby the reference lists of sources selected for full-text review were also examined for additional studies to include in the final review sample.

Study Selection

Eligibility Criteria
Studies on HIT and cancer screening before January 2015 are covered in prior publications [15,17,18]. Our search was designed to build on that work, so it was limited to studies published from 2015 to 2021 (the time at which we started the review process). Studies were considered eligible for inclusion if they (1) were US-based, reported in the English language, and published between January 2015 and June 2021; (2) reported on activities conducted in the primary care setting; (3) focused on evidence-based cancer screening; (4) involved the use of HIT to support this screening; (5) were related to specific workflow steps involved in conducting cancer screening in primary care (identifying patients due for screening at the point of care or in panel management, obtaining results of past screenings through data exchange, or providing appropriate follow-up care); and (6) targeted screening for breast, colorectal, or cervical cancer. A checklist of these criteria was created to guide the selection of relevant studies and then pilot-tested in a subsample of articles (n=60) and refined (COJ, RG, and RX) to ensure that its criteria could be applied consistently. The checklist was supported by a glossary of key terms to ensure shared understanding across reviewers of potential studies. The final checklist and glossary are provided in Multimedia Appendices 3 and 4, respectively. All study designs were eligible for inclusion as long as the study included some description of how HIT was used to support breast, colorectal, and cervical cancer screening in primary care settings. If a study was an evidence review (eg, systematic review or narrative review), only studies included in the final sample of the review and published between January 2015 and June 2021 were assessed for potential inclusion. If multiple publications described a single intervention but described different approaches for using HIT, all applicable studies were assessed for inclusion.

Dual Screening Review

Results of the search strategies described above were imported and managed in Zotero [35]. The first author (COJ) removed duplicate studies. Then, reviewers in eight 2-person teams were assigned studies to dual screen [36] (team 1: COJ and RG; team 2: AH and HA; team 3: LD and RX; team 4: KR and JMF; team 5: KHC and EB; team 6: KAR and IC; team 7: MMK and ATR; and team 8: MIF and DIA). Dual screening was performed in 2 steps. First, study titles and abstracts were dual screened by each review team using the inclusion and exclusion checklist to assess eligibility. Second, studies included for full-text dual screening were assessed by the same review teams for final inclusion in the scoping review. Any discrepancies that emerged within a review team were reconciled by consensus. The first
and senior authors (COJ and RG) provided final decisions for any studies that could not be reconciled by a review team.

Data Charting

A data charting form was developed using Qualtrics, a web-based survey software, to systematically extract information from studies selected for inclusion in the review analyses. The form was initially pilot-tested on 2 articles and refined (COJ, RG, and HA). Next, the review teams extracted information from their assigned studies. Extracted data elements included study citation, publication year, publication type, study design, study setting, sample composition by race or ethnicity, and cancer screening focus (breast, colorectal, and cervical cancer). Extracted characteristics of the relevant HIT tools involved in a given study included type, users, functions, purpose (intervention or implementation strategy supporting an intervention), and supported cancer screening activities. Data elements were extracted in multiple choice or free-text form, depending on the type of data. Multiple implementation frameworks [37-40] were used to guide data extraction. A check of at least 50% (49/101 studies) of extracted studies suggested that data charting quality was high and the agreement rate between the initial reviewers and the reviewers that conducted the quality check was >90%.

Multiple implementation frameworks [37-40] were used to guide data extraction. Specifically, the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework [37] guided the extraction of dissemination and implementation outcomes: target end users (clinical staff and patients) of HIT (Reach), HIT impact on cancer screening in primary care (Effectiveness), the rate of HIT adoption (Adoption), the extent to which a given HIT-based intervention was implemented (Implementation), and the extent to which sustainability of HIT adoption was measured (Maintenance). Assessment of the evidence on barriers and facilitators of HIT adoption was guided by the Integrated Technology Implementation Model (ITIM), which includes 12 inner and outer context concepts known to be central to the implementation and adoption of technology in health care settings, and is based on the Consolidated Framework for Implementation Research, adapted to HIT-based interventions [38]. Although technology frameworks have been used to investigate the usability and acceptance of HIT-based interventions [41-43], to our knowledge, the ITIM is the only model that incorporates theories of both implementation science and technology adoption. The Expert Recommendations for Implementing Change (ERIC) compilation [39] guided the categorization of discrete implementation strategies identified in the studies. The implementation strategies reporting the framework by Proctor et al [40] guided the extraction and analysis of implementation strategies used to support the HIT adoption.

Collating, Summarizing, and Reporting Results

Descriptive data were compiled and interpreted using Stata/MP (version 15.1; StataCorp LLC) to quantify the frequencies of extracted data in discrete fields. Free text data charted in Qualtrics were exported to Excel (Microsoft Corp) for qualitative content analysis [44,45]. Authors (COJ, JC, RX, and RG) reviewed and categorized free text for HIT characteristics, RE-AIM domains, implementation barriers, facilitators, and core elements of implementation strategies (eg, actor and target of action). Most analyses used an iterative process, which involved initial coding and identification of themes (ie, categories) by 2 reviewers, resolving discrepancies and refining categories through team discussion, and recoding the text using finalized categories. Multimedia Appendix 5 provides details about these procedures.

Consultation

Authors (RG, JC, RX, HA, and AH) were consulted at each stage of the scoping review to provide input on the search, data abstraction, and interpretation of the results. We also consulted with implementation science experts about the conceptual frameworks selected for this study.

Results

Literature Search

The search yielded an initial total of 618 studies (Figure 1). After removing duplicates, 485 titles and abstracts were assessed for inclusion. Among these, 350 studies were excluded as not meeting the inclusion criteria. Full-text review was conducted on 135 records that met the inclusion criteria. A snowball search yielded an additional 115 studies that were assessed for eligibility. A final total of 101 studies met the inclusion criteria. Multimedia Appendix 6 provides a complete list of these studies.
Characteristics of the Included Studies

Included studies were published between January 2015 and June 2021 (Table 1). Most studies were peer-reviewed (92/101, 91.1%). Study design was mostly nonexperimental (descriptive: 18/101, 17.8% or observational: 15/101, 14.9%) in comparison to experimental (randomized controlled trials: 29/101, 28.7%), quasi-experimental (pre-post design: 21/101, 20.8%; nonrandomized controlled trials: 5/101, 5%; or other quasi-experimental studies: 3/101, 3%), and other studies (10/101, 9.9%). Most studies covered HIT targeting colorectal cancer screening (83/101, 82.2%), followed by breast cancer screening (28/101, 27.7%) and cervical cancer screening (19/101, 18.8%); these sum up >101 as some addressed more than one type of cancer screening.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. HIT: health IT.
Table 1. Characteristics of the included studies (N=101).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Colorectal cancer (n=83), n (%)</th>
<th>Breast cancer (n=28), n (%)</th>
<th>Cervical cancer (n=19), n (%)</th>
<th>Total, (N=101), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publication year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>13 (15.7)</td>
<td>4 (14.3)</td>
<td>1 (5.3)</td>
<td>15 (14.9)</td>
</tr>
<tr>
<td>2016</td>
<td>15 (18.1)</td>
<td>6 (21.4)</td>
<td>3 (15.8)</td>
<td>18 (17.8)</td>
</tr>
<tr>
<td>2017</td>
<td>15 (18.1)</td>
<td>5 (17.9)</td>
<td>4 (21.1)</td>
<td>21 (20.8)</td>
</tr>
<tr>
<td>2018</td>
<td>16 (19.3)</td>
<td>2 (7.1)</td>
<td>3 (15.8)</td>
<td>18 (17.8)</td>
</tr>
<tr>
<td>2019</td>
<td>10 (12)</td>
<td>5 (17.9)</td>
<td>4 (21.1)</td>
<td>13 (12.9)</td>
</tr>
<tr>
<td>2020</td>
<td>9 (10.8)</td>
<td>5 (17.9)</td>
<td>3 (15.8)</td>
<td>11 (10.9)</td>
</tr>
<tr>
<td>2021</td>
<td>5 (6)</td>
<td>1 (3.6)</td>
<td>1 (5.3)</td>
<td>5 (5)</td>
</tr>
<tr>
<td><strong>Publication type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer-reviewed article</td>
<td>78 (94)</td>
<td>26 (92.9)</td>
<td>16 (84.2)</td>
<td>92 (91.1)</td>
</tr>
<tr>
<td>Report</td>
<td>1 (1.2)</td>
<td>2 (7.1)</td>
<td>2 (10.5)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Study protocol</td>
<td>3 (3.6)</td>
<td>—</td>
<td>—</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (1.2)</td>
<td>—</td>
<td>1 (5.3)</td>
<td>2 (2)</td>
</tr>
<tr>
<td><strong>Study design</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonexperimental</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Descriptive</td>
<td>15 (18.1)</td>
<td>5 (17.9)</td>
<td>4 (21.1)</td>
<td>18 (17.8)</td>
</tr>
<tr>
<td>Observational</td>
<td>11 (13.3)</td>
<td>9 (32.1)</td>
<td>4 (21.1)</td>
<td>15 (14.9)</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td>24 (28.9)</td>
<td>7 (25)</td>
<td>5 (26.3)</td>
<td>29 (28.7)</td>
</tr>
<tr>
<td>Quasi-experimental</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-post study design</td>
<td>17 (20.5)</td>
<td>5 (17.9)</td>
<td>2 (10.5)</td>
<td>21 (20.8)</td>
</tr>
<tr>
<td>Non-RCT</td>
<td>3 (3.6)</td>
<td>1 (3.6)</td>
<td>3 (15.8)</td>
<td>5 (5)</td>
</tr>
<tr>
<td>Other quasi-experimental</td>
<td>3 (3.6)</td>
<td>—</td>
<td>—</td>
<td>3 (3.0)</td>
</tr>
<tr>
<td>Other study designs</td>
<td>10 (12)</td>
<td>1 (3.6)</td>
<td>1 (5.3)</td>
<td>10 (9.9)</td>
</tr>
</tbody>
</table>

\(^a\)Percentages were calculated based on column totals.

\(^b\)Publication year represents studies published from January 2015 to June 2021.

\(^c\)Not available.

\(^d\)RCT: randomized controlled trial.

Characteristics of the primary care settings where the research in the included studies was conducted are shown in Table 2. Approximately half of the included studies (52/101, 51.5%) reported on practice location. Most studies involving colorectal (22/83, 27%) or breast (6/28, 21%) cancer screening were conducted in urban areas, and most studies on cervical cancer screening (5/19, 26%) were conducted in rural areas. Studies on colorectal cancer screening were primarily conducted in federally qualified health centers (20/83, 24%); most of those on breast and cervical cancer screening were conducted in academic-based clinics (9/28, 32% and 5/19, 26%, respectively). More than half (59/101, 58.4%) of the included studies (colorectal: 47/83, 57%; breast cancer: 17/28, 61%; and cervical: 8/19, 42%) reported information on racial or ethnic minoritized participants (patients from racial or ethnic minority groups). Of these 59 studies, 34 (58%) reported that ≤50% of study participants were members of racial or ethnic minority populations.
Table 2. Primary care practice characteristics of the included studies (N=101).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Colorectal cancer (n=83), n (%)</th>
<th>Breast cancer (n=28), n (%)</th>
<th>Cervical cancer (n=19), n (%)</th>
<th>Total, (N=101), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Practice location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>22 (26.5)</td>
<td>6 (21.4)</td>
<td>2 (10.5)</td>
<td>26 (25.7)</td>
</tr>
<tr>
<td>Rural</td>
<td>11 (13.3)</td>
<td>5 (17.9)</td>
<td>5 (26.3)</td>
<td>15 (14.9)</td>
</tr>
<tr>
<td>Combination of urban and rural</td>
<td>11 (13.3)</td>
<td>2 (7.1)</td>
<td>3 (15.8)</td>
<td>11 (10.9)</td>
</tr>
<tr>
<td>Not reported</td>
<td>39 (47)</td>
<td>15 (53.6)</td>
<td>9 (47.4)</td>
<td>49 (48.5)</td>
</tr>
<tr>
<td><strong>Practice type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic-based clinic</td>
<td>17 (20.5)</td>
<td>9 (32.1)</td>
<td>5 (26.3)</td>
<td>22 (21.8)</td>
</tr>
<tr>
<td>Federally Qualified Health Centers</td>
<td>20 (24.1)</td>
<td>1 (3.6)</td>
<td>3 (15.8)</td>
<td>21 (20.8)</td>
</tr>
<tr>
<td>Freestanding or other</td>
<td>18 (21.7)</td>
<td>4 (14.3)</td>
<td>3 (15.8)</td>
<td>21 (20.8)</td>
</tr>
<tr>
<td>Hospital-based clinic</td>
<td>10 (12)</td>
<td>2 (7.1)</td>
<td>1 (5.3)</td>
<td>12 (11.9)</td>
</tr>
<tr>
<td>Not reported</td>
<td>18 (21.7)</td>
<td>12 (42.9)</td>
<td>7 (36.8)</td>
<td>25 (24.8)</td>
</tr>
<tr>
<td><strong>Sample percentage of racial or ethnic minority groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50%</td>
<td>25 (30.1)</td>
<td>13 (46.4)</td>
<td>6 (31.6)</td>
<td>34 (33.7)</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>22 (26.5)</td>
<td>4 (14.3)</td>
<td>2 (10.5)</td>
<td>25 (24.8)</td>
</tr>
<tr>
<td>Not reported</td>
<td>36 (43.4)</td>
<td>11 (39.3)</td>
<td>11 (57.9)</td>
<td>42 (41.6)</td>
</tr>
</tbody>
</table>

aPercentages were calculated based on column totals.

**Characteristics of the HIT Interventions**

Our definitions of HIT tool types and functions and the types of cancer screening activities they supported are provided in Multimedia Appendix 4. Of the 101 included studies, 66 (65.3%) reported on interventions involving 1 HIT tool and 35 (34.7%) reported on interventions involving >1 HIT tool (Table 3). In these studies, the HIT tool was either the intervention of focus, one component of a multicomponent intervention that also included non-HIT elements, or was used as an implementation strategy to support the intervention of focus.

Most of the included studies (85/101, 84.2%) involved EHR-based HIT tools (Table 3). Web-based (18/101, 17.8%) and other types of HIT tools (19/101, 18.8%) were less common. The HIT function most commonly involved in included studies was clinical decision support (CDS) across all cancer screening types (Table 3). CDS tools for panel management were most common in studies involving colorectal cancer screening (50/83, 60%). CDS at the point of care was commonly used in studies on breast (16/28, 57%) and cervical cancer screening (12/19, 63%). Other commonly studied HIT functions included risk identification (colorectal: 13/83, 16% and cervical: 6/19, 32%), patient decision aids (colorectal: 13/83, 16% and breast: 9/28, 32%), and tools for tracking patient adherence to recommended care (colorectal: 27/83, 33% and cervical: 6/19, 32%).

The cancer screening activities were primarily related to identifying patients for screening in panel management (colorectal: 50/83, 60%; breast: 8/28, 29%; and cervical: 7/19, 37%) and at the point of care (colorectal: 39/83, 47%; breast: 15/28, 54%; and cervical: 12/19, 63%). Other commonly supported cancer screening activities included follow-up care for referral (colorectal: 36/83, 43%; breast: 7/28, 25%; and cervical: 7/19, 37%) and for positive or abnormal screening results (colorectal: 12/83, 15% and cervical: 5/19, 26%; Table 3).
Table 3. Characteristics of the health IT (HIT) sources and functions used to promote cancer screening in primary care, as represented in the included studies (N=101).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Colorectal cancer (n=83), n (%)</th>
<th>Breast cancer (n=28), n (%)</th>
<th>Cervical cancer (n=19), n (%)</th>
<th>Total (N=101), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Using single or multiple HIT tools</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single HIT tools</td>
<td>53 (63.9)</td>
<td>22 (78.6)</td>
<td>14 (73.7)</td>
<td>66 (65.3)</td>
</tr>
<tr>
<td>Multiple HIT tools</td>
<td>30 (36.1)</td>
<td>6 (21.4)</td>
<td>5 (26.3)</td>
<td>35 (34.7)</td>
</tr>
<tr>
<td><strong>HIT sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EHR&lt;sup&gt;b&lt;/sup&gt; based</td>
<td>74 (89.2)</td>
<td>20 (71.4)</td>
<td>18 (94.7)</td>
<td>85 (84.2)</td>
</tr>
<tr>
<td>Web based</td>
<td>11 (13.3)</td>
<td>9 (32.1)</td>
<td>3 (15.8)</td>
<td>18 (17.8)</td>
</tr>
<tr>
<td>Other or unclear</td>
<td>15 (18.1)</td>
<td>3 (10.7)</td>
<td>3 (15.8)</td>
<td>19 (18.8)</td>
</tr>
<tr>
<td><strong>HIT functions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDS&lt;sup&gt;c&lt;/sup&gt; panel management or outreach</td>
<td>50 (60.2)</td>
<td>7 (25)</td>
<td>9 (47.4)</td>
<td>57 (56.4)</td>
</tr>
<tr>
<td>CDS point of care</td>
<td>41 (49.4)</td>
<td>16 (57.1)</td>
<td>12 (63.2)</td>
<td>48 (47.5)</td>
</tr>
<tr>
<td>Risk identification</td>
<td>13 (15.7)</td>
<td>5 (17.9)</td>
<td>6 (31.6)</td>
<td>18 (17.8)</td>
</tr>
<tr>
<td>Patient decision aid</td>
<td>13 (15.7)</td>
<td>9 (32.1)</td>
<td>2 (10.5)</td>
<td>18 (17.8)</td>
</tr>
<tr>
<td>Provider assessment and feedback</td>
<td>11 (13.3)</td>
<td>1 (3.6)</td>
<td>1 (5.3)</td>
<td>12 (11.9)</td>
</tr>
<tr>
<td>Tracking patient adherence</td>
<td>27 (32.5)</td>
<td>4 (14.3)</td>
<td>6 (31.6)</td>
<td>30 (29.7)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (3.6)</td>
<td>—</td>
<td>—</td>
<td>3 (3.0)</td>
</tr>
<tr>
<td><strong>Cancer screening activities supported by HIT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel management</td>
<td>50 (60.2)</td>
<td>8 (28.9)</td>
<td>7 (36.8)</td>
<td>56 (55.4)</td>
</tr>
<tr>
<td>Point of care</td>
<td>39 (47)</td>
<td>15 (53.6)</td>
<td>12 (63.2)</td>
<td>45 (44.6)</td>
</tr>
<tr>
<td>Follow-up (referral)</td>
<td>36 (43.4)</td>
<td>7 (25.0)</td>
<td>7 (36.8)</td>
<td>41 (40.6)</td>
</tr>
<tr>
<td>Follow-up (abnormal or positive result)</td>
<td>12 (14.5)</td>
<td>2 (7.1)</td>
<td>5 (26.3)</td>
<td>17 (16.8)</td>
</tr>
<tr>
<td>Acquire previous results</td>
<td>7 (8.4)</td>
<td>2 (7.1)</td>
<td>4 (21.1)</td>
<td>10 (9.9)</td>
</tr>
<tr>
<td>Other</td>
<td>21 (25.3)</td>
<td>11 (39.3)</td>
<td>5 (26.3)</td>
<td>24 (23.8)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Percentages were calculated based on column totals. Some studies featured >1 HIT source, function, and cancer screening activity. As a result, these categories are not mutually exclusive and will not necessarily sum to 100%. Refer to Multimedia Appendix 4 for definitions of the terms used in this table.

<sup>b</sup>EHR: electronic health record.

<sup>c</sup>CDS: clinical decision support.

<sup>d</sup>Not available.

**Reporting on RE-AIM Outcomes**

**Overview**

A summary of reporting on RE-AIM outcomes is provided in Table 4.
Table 4. Reporting on Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) outcomes for health IT (HIT) targeting cancer screening in primary care.

<table>
<thead>
<tr>
<th>RE-AIM domains</th>
<th>Data charted</th>
<th>Cancer screening type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Colorectal cancer</td>
</tr>
<tr>
<td>Reach</td>
<td>Was the number of targeted staff or patients for HIT-based intervention reported</td>
<td>High(^a)</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Did the HIT tools show positive results in the cancer screening intervention</td>
<td>High</td>
</tr>
<tr>
<td>Adoption</td>
<td>Rate of HIT adoption</td>
<td>Low</td>
</tr>
<tr>
<td>Implementation</td>
<td>Barriers, facilitators, and implementation strategies used related to HIT</td>
<td>Moderate</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Was the sustainment of HIT adoption measured</td>
<td>Low</td>
</tr>
</tbody>
</table>

\(^a\)Low: <25% of the included studies for each cancer screening type category, moderate: 25% to 50% of the included studies for each cancer screening type category, and high: >50% of the included studies for each cancer screening type category. Percentages were calculated with respect to the included studies for each cancer screening type category.

**Effectiveness**

Of the 101 included studies, 24 (23.8%) reported on the effectiveness of HIT targeting breast (14/28, 50% of breast cancer–relevant studies) and cervical cancer screening (10/19, 53% of cervical cancer–relevant studies; Multimedia Appendix 7 includes a table with these results). Of the 101 included studies, 54 (53.5%) reported the effectiveness of HIT targeting colorectal cancer screening (54/83, 65% of colorectal cancer–relevant studies). Among studies reporting on effectiveness, most-reported positive outcomes (improved screening rate) associated with the use of HIT (36/54, 67% of colorectal cancer–relevant studies; 9/14, 64% of breast cancer–relevant studies; and 7/10, 70% of cervical cancer–relevant studies). This evidence mostly represented CDS used during panel management (22/83, 27% of colorectal cancer–relevant studies) or at the point of care (5/28, 18% of breast cancer–relevant studies and 5/19, 26% of cervical cancer–relevant studies; Multimedia Appendix 7).

**Reach, Adoption, and Maintenance**

Among the 101 included studies, 79 (78.2%) reported on the reach of HIT-based interventions. Most of the studies focused on reach involved HIT for colorectal cancer screening (63/83, 76% of colorectal cancer–relevant studies studies). The reach of HIT-based interventions targeting breast cancer screening was reported in 82% (23/28) of the breast cancer–relevant studies and in 74% (14/19) of the studies targeting cervical cancer screening. Overall, 15.8% (16/101) of the studies reported on HIT adoption (colorectal: 10/83, 12%; breast: 9/28, 32%; and cervical: 1/19, 5%), and 2% (2/101) of the studies reported on maintenance of HIT-based interventions. Of those that reported on adoption, there was mostly a low rate of adoption (≤50%) across all cancer screening types (Multimedia Appendix 8 includes a table with these results).

**Implementation**

The proportion of studies reporting on the implementation of the HIT ranged from 25% to 50% for those related to colorectal cancer screening (Table 4). It was reported in <25% of the studies related to HIT targeting breast and cervical cancer screening. Implementation barriers, facilitators, and strategies related to HIT adoption across all cancer screening types are described further in the next 2 sections.

**Implementation Barriers and Facilitators of HIT Adoption**

A total of 34 studies reported on barriers and 37 studies reported on facilitators to implementing the HIT-based interventions of focus in primary care (Table 5). The most-reported barriers and facilitators were related to the ITIM constructs inner context (barriers: 17/34, 50% and facilitators: 14/37, 38%), nature of the innovation (barriers: 15/34, 44% and facilitators: 17/37, 46%), and outer context (barriers: 11/34, 32% and facilitators: 9/37, 24%). Inner context barriers included limited staff time to use the HIT and adoption competing with other clinic priorities. Inner context facilitators included having dedicated staff assigned to operate and manage a given HIT tool, and organizational policies supporting HIT adoption. Barriers related to the nature of the innovation included inaccurate cancer screening data reported by the HIT intervention and the burden of HIT development and maintenance. Facilitators related to the nature of the innovation included that HIT automation and customization features reduced staff resources and time needed in providing care. Outer context barriers included challenges involved with working with an EHR vendor to activate and update the tool and challenges with accessing screening results conducted outside the clinics. Outer context facilitators included Medicaid expansion including cancer screening as an incentivized metric and the clinic being a Federally Qualified Health Center, which necessitated responsiveness to such metrics. A table with more examples of barriers and facilitators is provided in Multimedia Appendix 9.
Table 5. Reporting on the barriers and facilitators of health IT adoption aligned with the Integrated Technology Implementation Model (ITIM).

<table>
<thead>
<tr>
<th>ITIM constructs</th>
<th>Barriers (n=34), n (%)</th>
<th>Facilitators (n=37), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adoption or adopters</td>
<td>2 (6)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Communication</td>
<td>6 (18)</td>
<td>5 (14)</td>
</tr>
<tr>
<td>Economic environment</td>
<td>5 (15)</td>
<td>6 (16)</td>
</tr>
<tr>
<td>Facilitators (boundary spanner)</td>
<td>_</td>
<td>4 (11)</td>
</tr>
<tr>
<td>Implementation</td>
<td>3 (9)</td>
<td>9 (24)</td>
</tr>
<tr>
<td>Inner context</td>
<td>17 (50)</td>
<td>14 (38)</td>
</tr>
<tr>
<td>Interfacing systems</td>
<td>5 (15)</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Leadership</td>
<td>2 (6)</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Nature of the innovation</td>
<td>15 (44)</td>
<td>17 (46)</td>
</tr>
<tr>
<td>Outer context</td>
<td>11 (32)</td>
<td>9 (24)</td>
</tr>
<tr>
<td>Users (adopters)</td>
<td>9 (26)</td>
<td>4 (11)</td>
</tr>
<tr>
<td>Workflow</td>
<td>9 (26)</td>
<td>11 (30)</td>
</tr>
</tbody>
</table>

aPercentages were calculated with respect to the total studies that reported barriers or facilitators. Some studies featured both barriers and facilitators to health IT adoption for cancer screening in primary care. As a result, these categories are not mutually exclusive and will not necessarily sum to 100%.

bNot available.

Implementation Strategies to Support HIT Adoption

Implementation strategies targeting HIT adoption were reported in 24% (24/101) of the included studies. Those reported were mapped to 22 implementation strategies from the ERIC compilation [39] (Multimedia Appendix 10). Of the studies reporting implementation strategies, >50% used ≥2 strategies and >50% reported strategies promoting HIT use for colorectal cancer screening. Common strategies to promote HIT use among all cancer screening types included central technical assistance, conducting small tests of change, and educational meetings. A table with more examples is available in Multimedia Appendix 10. Reported evidence mapped to the domains formulated by Proctor et al [40] (Table 6) and were mostly focused on describing implementation strategies to support HIT adoption for colorectal cancer screening (22/83, 27% of colorectal cancer–relevant studies) in comparison to breast (6/28, 21% of breast cancer–relevant studies) and cervical cancer screening (4/19, 21% of cervical cancer–relevant studies). Overall, less than half of the included studies, for each cancer screening type, reported evidence in accordance with each implementation strategy domain.

Table 6. Reporting on the implementation strategies used to support health IT adoption.

<table>
<thead>
<tr>
<th>Implementation strategy domains by Proctor et al [40]</th>
<th>Data charted</th>
<th>Cancer screening type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actor</td>
<td>Who delivers the strategy</td>
<td>Colorectal cancer</td>
</tr>
<tr>
<td>Action</td>
<td>Procedures to conduct the strategy</td>
<td>Moderate&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Target of action</td>
<td>Intent of action</td>
<td>Low</td>
</tr>
<tr>
<td>Temporality</td>
<td>When does the strategy happen</td>
<td>Low</td>
</tr>
<tr>
<td>Dose</td>
<td>Frequency or intensity</td>
<td>Low</td>
</tr>
<tr>
<td>Implementation outcomes affected</td>
<td>What will the strategy change</td>
<td>Low</td>
</tr>
<tr>
<td>Justification</td>
<td>Purpose of the strategy</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

<sup>a</sup>Low: <25% of the included studies for each cancer screening type category, moderate: 25% to 50% of the included studies for each cancer screening type category, and high: >50% of the included studies for each cancer screening type category.

Discussion

Principal Findings

This scoping review summarizes the state of the science about the implementation of HIT-based interventions targeting breast, cervical, or colorectal cancer screening in primary care. Previous reviews identified the positive impact of HIT-based interventions throughout the cancer care continuum, including cancer screening [15,17,18]. This review adds to prior evidence by bringing an implementation science perspective; this is needed because the impact of HIT-based interventions is limited by the extent to which such interventions are effectively integrated into practice. This scoping review provides updated evidence
up to 2021. This is not a systematic review; our goal was to identify knowledge gaps. Results indicate that key knowledge gaps related to the implementation of HIT in cancer screening in primary care include (1) the effectiveness of HIT targeting breast and cervical cancer screening, (2) HIT adoption in diverse primary care settings, (3) the implementation strategies that support the adoption of HIT, and (4) equitable reach or adoption of HIT. Addressing these evidence gaps may be critical to supporting the implementation of high-quality primary care [46].

Knowledge Gap 1: Limited Evidence on the Effectiveness of HIT Targeting Breast and Cervical Cancer Screening

This review emphasizes the need to improve the evidence on HIT effectiveness, especially HIT targeting breast and cervical cancer screening uptake, in diverse primary care settings. Effectiveness outcomes included, but were not limited to, improvements in cancer screening initiation by the patient or provider and patient completion of cancer screening. Although the use of HIT-based interventions was associated with improved screening outcomes for all 3 cancer types, there were far fewer studies of HIT effectiveness for breast and cervical cancer screening (a combined total of 24 studies) in comparison to the 54 studies involving colorectal cancer screening. Furthermore, most studies related to HIT targeting breast or cervical cancer prevention were conducted in academic medical centers and were not readily generalizable to other primary care settings. This limited evidence is concerning as both are common cancers, and evidence-based guidelines for such screenings are not met in many patient populations.

In addition, the lack of reporting on HIT effectiveness was especially common in studies in which HIT was part of a multicomponent intervention [44]; thus, even if the effectiveness of the overall intervention was reported, the impact of the HIT element of the intervention was not clear. More research is needed to establish the effectiveness of HIT targeting cancer screening in diverse primary care settings, including trials of the individual and combined effect of HIT within multicomponent interventions. The need for an improved understanding of the effectiveness of HIT is especially salient given that national programs (eg, Promoting Interoperability Program, formerly Meaningful Use) promote the use of HIT in health care settings [47] as a means to improve health outcomes.

Knowledge Gap 2: Limited Evidence of the Reach, Adoption, and Maintenance of Effective HIT Targeting Cancer Screening

The limited reporting on the reach, adoption, and maintenance of such interventions aligns with the known lack of reporting on these implementation outcomes in analyses of other interventions [48]; the need to improve such reporting is well known in implementation science. When HIT adoption is not reported, it is difficult to assess an intervention’s population-level impact. In particular, if a limited number of potential users adopt an intervention, even when there is good reach and it is highly effective, population-level impacts may be low. Where adoption was reported, its rates were generally low (≤50%), underscoring the need for further research on improving the uptake of effective HIT [49]. When implementation barriers and facilitators to HIT adoption were reported, most related to inner context, outer context, and the nature of the innovation (including a given HIT tool’s function). Future research should assess which combination of these contextual factors is associated with the adoption of HIT with varied functions when used in different workflow steps (ie, panel management, point of care, and follow-up care). To further understand how contextual factors impact care teams’ adoption of HIT for cancer screening, there is also a need for more widespread reporting on practice type, which was rarely noted in the studies included here. Similarly, few studies reported on the sustainment of tool adoption. This evidence gap is seen throughout the implementation science literature [50]; improved knowledge of how to sustain the use of effective interventions is critical to maximizing their impact. Knowledge gap 3 describes the need for evidence on how to improve the adoption and maintenance of HIT-based interventions targeting cancer screening in primary care. We also posit that the lack of evidence on such interventions’ reach is relevant to how such interventions support equity in cancer screening, as discussed in knowledge gap 4.

Knowledge Gap 3: Limited Evidence on Implementation Strategies That Support the Adoption of HIT Targeting Cancer Screening in Primary Care

A total of 24 studies (<25% of the included studies) reported on strategies used to support the adoption of HIT-based interventions, and few of these assessed the effectiveness of the strategies. This is complicated by the fact that in some cases a given HIT tool was considered the intervention or an intervention component, and in others, it was considered an implementation strategy for supporting the adoption of a clinical intervention. In the implementation science literature, the boundaries between clinical intervention and implementation intervention and between implementation intervention and implementation strategies are not always clear, adding complexity to this reporting.

Research is needed on how to support the adoption of HIT-based interventions targeting cancer screening using implementation strategies, how to use HIT as an implementation strategy, and what types of support strategies are used even in reports on HIT-based interventions’ impact. Reporting must strive to clearly differentiate between these approaches; the need for better reporting on implementation strategies is well known [51-53]. Although such reporting can be resource intensive, methods are emerging to facilitate it [40,54].

Research is also needed to specify how effectively different implementation strategies support the adoption of different HIT-based interventions in different care settings. Known barriers and facilitators to HIT adoption, in general, may also be impactful for HIT targeting cancer screening. For example, evidence indicates that barriers to HIT use include inadequate training for care teams on using EHR functions to their full potential [55-58]. Thus, effective implementation strategies for HIT targeting cancer prevention may involve training.
Knowledge Gap 4: Limited Evidence on the Reach and Equitable Implementation of HIT for Cancer Screening in Primary Care

The equitable reach of HIT tools for cancer screening is poorly described. A few studies specifically focused on racial or ethnic minority groups; many were conducted in federally qualified health centers, which often serve racial or ethnic minority groups. Relevant data were reported in just 58.4% (59/101) of the studies included here. However, where such data were reported, eligible patients reached by the HIT interventions had a lower percentage of non-White patients than would be expected for the populations served, suggesting inequities in reach or underreporting. This is concerning, as racial disparities in cancer screening persist [59-61], and previous research found that interventions targeting breast or cervical cancer screening are less likely to target patients considered most at risk, for example, those in socioeconomically and racial or ethnic minoritized groups [5]. Findings from this scoping review underscore the need to understand potential drivers of these inequities (eg, design flaws in algorithms used to identify eligible patients and clinician bias in applying the HIT tool) and solutions to mitigate these inequities. One step toward addressing this inequity must involve improving reporting on how HIT is used in diverse patient populations. The well-documented need to improve reporting of race or ethnicity in health care [62] likely exacerbates the lack of reporting on the comparative reach of the tools included in this review among different groups. Another step toward equitable reach of HIT is understanding and addressing barriers to the inclusion of racial or ethnic minoritized patients in research on HIT adoption and impact. Future research on HIT adoption for cancer screening should explore strategies that support documentation, recruitment, and retention of racial or ethnic minoritized patients [63].

Limitations

HIT-based interventions might be used to improve outcomes at each step of the cancer control continuum, such as risk assessment, prevention, detection, diagnosis, treatment, survivorship, and end-of-life care [15]. This review was limited to cancer screening. Furthermore, although breast, colorectal, and cervical cancer are highly prevalent cancers whose detection is in the purview of primary care, no other cancers recommended for screening in primary care (eg, lung cancer) were included; future research could assess whether the gaps identified in this study are seen for a broader set of cancers. This review was limited to US studies; therefore, the relevance of the findings is limited to the context of HIT policies and infrastructure as applicable to US primary care settings. Another potential limitation is that urban or rural status was defined based on what each study reported, and they may have used different methods for making this characterization.

In addition, the overlapping quality of some HIT characteristic categories (tool types and functions) made it difficult to execute related data charting. Similarly, content analysis of HIT functions was complicated when implementation strategies overlapped or when studies did not specify which cancers were targeted by the strategies. Our definition of effectiveness did not capture screening outcomes related to each clinical workflow (eg, an intervention using CDS for panel management showed improvements in colorectal cancer screening but did not clarify how improvements impacted screening initiation, completion, or follow-up care). Finally, we followed the PRISMA-ScR guidelines [34] to examine a broad array of literature to include studies that are heterogeneous in design and quality [64]. Although our search strategy followed an iterative process, it is possible that some relevant existing articles were not captured; we sought to mitigate this using a snowball search.

Conclusions

In what is, to our knowledge, the first scoping review of the implementation of HIT-based interventions for cancer screening in primary care settings, we identified critical knowledge gaps. Little is known about the effectiveness of HIT-based interventions specifically targeting guideline-concordant breast and cervical cancer screening. Clarity is needed on the individual and combined effectiveness of HIT when integrated into a multicomponent intervention targeting cancer screening. Even less is known about how to enhance the adoption of cancer-targeted HIT in primary care. The potential for inequities in the reach of HIT for cancer screening remains underexplored. Research is necessary on implementation strategies to promote equitable access, ensuring that the potential benefits of HIT for population health are realized across diverse patient populations.

Acknowledgments

The authors would like to extend their gratitude to the Consortium for Cancer Implementation Science (CCIS). This research was conducted as part of the National Cancer Institute’s CCIS Technology in Implementation Science Action Group. This research was supported by the National Cancer Institute of the National Institutes of Health (P50CA244289, P50CA244693, P50CA244688, P50CA244690, P50CA244431), through funding provided as part of the Cancer Moonshot. This research was also supported by a doctoral scholarship award from the Provost Office of George Mason University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors’ Contributions

COJ and RG conceptualized the research questions. COJ conceptualized and led the data collection methods and analysis. COJ conducted a 3-step search process to identify relevant US-based peer-reviewed and gray literature studies. To ensure that the search yielded relevant studies, variations of the search strategy were pilot-tested by 3 authors (COJ, RG, and RX) and refined before the final search was conducted. COJ and RG conceptualized the eligibility criteria, and COJ, RG, and RX participated in

https://cancer.jmir.org/2024/1/e49002
the pilot-testing of the eligibility criteria before study selection. All authors conducted dual screening and applied the eligibility
criteria for study selection. COJ and RG developed the data charting form, and COJ, RG, and HA participated in pilot-testing the
data charting form before data charting. All authors conducted data charting of the final sample of studies selected for inclusion.
COJ, JC, RX, and RG synthesized the data and performed quality checks on the reported results. COJ led the manuscript
development. All coauthors had the opportunity to read, edit, and approve the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Database search strategy.
[DOCX File, 34 KB - cancer_v10i1e49002_app1.docx ]

Multimedia Appendix 2
Gray literature search strategy.
[DOCX File, 35 KB - cancer_v10i1e49002_app2.docx ]

Multimedia Appendix 3
Inclusion and exclusion criteria checklist.
[DOCX File, 34 KB - cancer_v10i1e49002_app3.docx ]

Multimedia Appendix 4
Glossary of key terms.
[DOCX File, 38 KB - cancer_v10i1e49002_app4.docx ]

Multimedia Appendix 5
Stage 5 procedures for qualitative content analysis.
[DOCX File, 33 KB - cancer_v10i1e49002_app5.docx ]

Multimedia Appendix 6
References to the included studies (N=101).
[DOCX File, 55 KB - cancer_v10i1e49002_app6.docx ]

Multimedia Appendix 7
Reporting status of health IT (HIT) effectiveness by cancer screening activities and HIT functions as represented in the included
studies.
[DOCX File, 41 KB - cancer_v10i1e49002_app7.docx ]

Multimedia Appendix 8
Reporting of health IT adoption as represented in the included studies.
[DOCX File, 35 KB - cancer_v10i1e49002_app8.docx ]

Multimedia Appendix 9
Barriers and facilitators (with examples) of health IT adoption aligned with the Integrated Technology Implementation Model.
[DOCX File, 39 KB - cancer_v10i1e49002_app9.docx ]

Multimedia Appendix 10
Expert Recommendations for Implementing Change implementation strategies used to support health IT adoption (n=24 studies).
[DOCX File, 46 KB - cancer_v10i1e49002_app10.docx ]

Multimedia Appendix 11
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 81 KB - cancer_v10i1e49002_app11.pdf ]

References


Abbreviations

CCIS: Consortium for Cancer Implementation Science

CDS: clinical decision support

EHR: electronic health record
Pediatric Cancer Communication on Twitter: Natural Language Processing and Qualitative Content Analysis

Nancy Lau1,2, PhD; Xin Zhao2, PhD; Alison O’Daffer3,4, BA; Hannah Weissman5; Krysta Barton6, PhD

1Center for Child Health, Behavior and Development, Seattle Children’s Research Institute, Seattle, WA, United States
2Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, United States
3Department of Psychiatry, University of California, San Diego, San Diego, CA, United States
4Center for Empathy and Technology, Sanford Institute for Empathy and Compassion, University of California, San Diego, San Diego, CA, United States
5Department of Psychology, Vanderbilt University, Nashville, TN, United States
6Biostatistics Epidemiology and Analytics for Research (BEAR) Core, Seattle Children’s Research Institute, Seattle, WA, United States

Corresponding Author:
Nancy Lau, PhD
Center for Child Health, Behavior and Development
Seattle Children’s Research Institute
1920 Terry Avenue
Seattle, WA, 98101
United States
Phone: 1 2068840569
Email: nancy.lau@seattlechildrens.org

Abstract

Background: During the COVID-19 pandemic, Twitter (recently rebranded as “X”) was the most widely used social media platform with over 2 million cancer-related tweets. The increasing use of social media among patients and family members, providers, and organizations has allowed for novel methods of studying cancer communication.

Objective: This study aimed to examine pediatric cancer–related tweets to capture the experiences of patients and survivors of cancer, their caregivers, medical providers, and other stakeholders. We assessed the public sentiment and content of tweets related to pediatric cancer over a time period representative of the COVID-19 pandemic.

Methods: All English-language tweets related to pediatric cancer posted from December 11, 2019, to May 7, 2022, globally, were obtained using the Twitter application programming interface. Sentiment analyses were computed based on Bing, AFINN, and NRC lexicons. We conducted a supplemental nonlexicon-based sentiment analysis with ChatGPT (version 3.0) to validate our findings with a random subset of 150 tweets. We conducted a qualitative content analysis to manually code the content of a random subset of 800 tweets.

Results: A total of 161,135 unique tweets related to pediatric cancer were identified. Sentiment analyses showed that there were more positive words than negative words. Via the Bing lexicon, the most common positive words were support, love, amazing, heaven, and happy, and the most common negative words were grief, risk, hard, abuse, and miss. Via the NRC lexicon, most tweets were categorized under sentiment types of positive, trust, and joy. Overall positive sentiment was consistent across lexicons and confirmed with supplemental ChatGPT (version 3.0) analysis. Percent agreement between raters for qualitative coding was 91%, and the top 10 codes were awareness, personal experiences, research, caregiver experiences, patient experiences, policy and the law, treatment, end of life, pharmaceuticals and drugs, and survivorship. Qualitative content analysis showed that Twitter users commonly used the social media platform to promote public awareness of pediatric cancer and to share personal experiences with pediatric cancer from the perspective of patients or survivors and their caregivers. Twitter was frequently used for health knowledge dissemination of research findings and federal policies that support treatment and affordable medical care.

Conclusions: Twitter may serve as an effective means for researchers to examine pediatric cancer communication and public sentiment around the globe. Despite the public mental health crisis during the COVID-19 pandemic, overall sentiments of pediatric cancer–related tweets were positive. Content of pediatric cancer tweets focused on health and treatment information, social support, and raising awareness of pediatric cancer.

(JMIR Cancer 2024;10:e52061) doi:10.2196/52061
Introduction

Social media platforms are widely used to exchange information and share resources. One such platform is Twitter (recently rebranded as “X”), a microblogging site with approximately 400 million global users. Social media platforms such as Twitter have been used by patients with health conditions, their caregivers, and other family members to connect with individuals in similar situations and learn from patients, researchers, and organizations worldwide. Patients with cancer, survivors of cancer, and their family members commonly use Twitter as a resource for treatment information and social support. Twitter users consist of a variety of cancer stakeholders including cancer centers, pharmaceutical companies, nonprofit organizations, medical providers, patients, and patients’ family and friends [1]. Individuals and organizations also use Twitter and other social media platforms to increase awareness and reach of cancer-related messages [2]. The increasing use of social media among patients, providers, and organizations has allowed for novel ways of studying cancer communication [3].

The global COVID-19 pandemic led to major changes in lifestyle, social distancing, and isolation that uniquely affected patients with cancer, caregivers, and other stakeholders. They were negatively impacted by overly taxed health care infrastructure and medical systems, restricted access to medical care, and a mental health crisis. Research on cancer during the pandemic spanned a range of topics including the global impact of COVID-19 on cancer care management [4,5]. Cancer survivors’ stressors during the pandemic included anxiety about in-person appointments, fear of cancer recurrence, medical care delays, uncertainty about future medical care, untreated symptoms, and mental health concerns [6]. Caregivers of patients with pediatric cancer experienced changes to their child’s medical care, financial disruptions, and emotional stress related to COVID-19 [7]. The COVID-19 pandemic was also associated with an increased risk of depression and loneliness in people living with cancer [8].

Twitter has been the most frequently used social media platform for public health surveillance since 2006, and there were over 2 million cancer-related tweets during the pandemic [9,10]. Previous studies have examined changes in public sentiment and the increasing use of Twitter during the pandemic [11,12]. For example, an analysis of Twitter showed that patients with cancer expressed significant negative sentiment during the COVID-19 pandemic [9]. Recent studies have examined the content of cancer-related tweets for different types of cancer diagnoses including lung, breast, and prostate cancer [13-15]. We only identified 2 studies thus far that have examined the pediatric cancer experience on Twitter. The first was a cross-sectional study examining the use of Twitter to discuss childhood cancer during Childhood Cancer Awareness Month [16]. The second used Twitter data to conduct a lexicon-based sentiment analysis of patients with pediatric cancer using the hashtag #ChildhoodCancer and found generally positive sentiment scores [17].

Lexicon-based sentiment analytic approaches determine positive and negative sentiments based on individual words [18]. Recent behavioral health studies have used lexicon-based approaches to analyze short text messages on social media platforms such as Twitter [19-23]. In addition to determining the positive or negative sentiment of words, the “NRC” lexicon assigns a sentiment type using the following additional emotion categories: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust [18]. Such analyses provide population-level insights into patterns of health information sharing and support-seeking on social media, and can inform the dissemination of time-critical information and resources during challenging times such as the COVID-19 pandemic. Additionally, the launch of Open AI’s chatbot, ChatGPT, provides a novel tool for nonlexicon-based sentiment analysis via text-based chat inquiries. Emerging research suggests that ChatGPT demonstrates superior performance in sentiment analysis of free-text responses [24].

In this study, we examined cancer-related tweets for pediatric cancer globally to capture the experiences of patients and survivors of cancer, their caregivers, medical providers, and other stakeholders. The objectives of our analysis using a Twitter data set over a time period representative of the impact of the COVID-19 pandemic were two-fold: (1) to quantitatively analyze the public sentiment of tweets related to pediatric cancer via lexicon-based and nonlexicon-based sentiment analytic approaches; and (2) to qualitatively examine the topics relevant to cancer diagnosis, treatment, and survivorship covered with hashtags commonly associated with pediatric cancer via a directed content analysis.

Methods

Ethical Considerations

This study did not require institutional review board approval because we used publicly available social media data that do not involve human subjects and do not fall within the scope of Human Subjects Research. Seattle Children’s Hospital’s Institutional Review Board uses the US Department of Health and Human Services (DHHS) definition of Human Subjects Research. Human Subjects Research under DHHS regulations is defined as the investigator obtaining information through intervention or interaction with living individuals; or obtaining, using, studying, analyzing, or generating identifiable private information. Our research is Nonhuman Subjects research according to the Seattle Children’s HRP-101 Human Research Protection Program Plan, P. 3, which uses the DHHS definition of Human Subjects Research [25,26]. The publicly available social media data reported in this paper have been anonymized and contains no IDs, user names, or nonparaphrased tweets.
Data Collection
In this study, we examined pediatric cancer–related communication on Twitter encompassing a representative timeframe ranging from before to after the COVID-19 pandemic. We obtained a total of 182,628 publicly available global tweets from December 11, 2019, to May 7, 2022, using the Twitter application programming interface. An example of the query and timeline information using “#teenagecancer” is available in Multimedia Appendix 1. For this study, we restricted our collection to English-only tweets. We identified a list of hashtags commonly associated with pediatric cancer: #childhoodcancer, #childhoodcancerawareness, #childhoodcancerday, #internationalchildhoodcancerday, #kidsgetcancertoo, #pediatriccancer, #pediatriconcology, #teenagecancer. These 8 keywords were selected because they were representative of hashtags frequently used for pediatric cancer. The prepandemic period was designated as December 2019 to February 2020. The pandemic time period was designated as March 2020 to June 2020. Lockdown and mandatory stay-at-home orders were issued in 42 US states and territories across the United States between March and May 2020 during the height of the pandemic [27]. The postpandemic time period was designated as July 2020 to May 2022 after mandatory stay-at-home orders were lifted in all states across the United States. Removing duplicates resulted in a total of 161,135 tweets from 40,289 unique users. All unique tweets were used for lexicon-based sentiment analysis. Among the 161,135 tweets from 40,289 unique accounts, we then randomly sampled a subset of 800 tweets and analyzed them using a directed content approach. Of the subset of tweets, 300 were randomly sampled and proportionately stratified by pandemic period (prepandemic, during the pandemic, and postpandemic).

Sentiment Analysis
Overview
“Sentiment analysis” or “opinion mining” is a natural language processing technique used to analyze and extract insights from text data, enabling the identification and understanding of the sentiment, emotions, and subjective opinions expressed within the text, which can be valuable for various applications such as market research, customer feedback analysis, and social media monitoring. We used lexicon-based approaches to conduct analyses using the full data set of tweets. Nonlexicon-based approaches can be used to evaluate whether the results may align with lexicon-based analysis. Thus, we used ChatGPT to conduct supplemental analyses on a randomly selected subsample of tweets.

Lexicon-Based Approaches
All data preprocessing, cleaning, and analyses were performed in R (version 4.2.2; R Foundation for Statistical Computing). We used “saotd” for preprocessing and initial analyses [28]. Nonlanguage elements such as symbols, weblinks, punctuation, emojis, and stop words, such as “the” and “of,” were removed. Sentiment scores were first computed based on the Bing lexicon, and we presented the most common positive and negative words within the data set. Additional analyses were conducted using the “syuzhet” package. We computed sentiments based on “Bing,” “AFINN,” and “NRC” lexicons. The “Bing” lexicon was developed by Liu [29] and categorizes 6788 English words into positive and negative categories. The “NRC” lexicon includes 6468 English words and classifies words as positive or negative sentiments and includes emotional categories of anger, anticipation, disgust, fear, joy, sadness, surprise, and trust [30]. The “AFINN” lexicon includes 2476 English words that were labeled with a value between –5 (negative sentiment) and +5 (positive sentiment) [31]. We used the “get_sentiment” function in the “syuzhet” package to calculate sentiment scores. Final sentiment scores were generated for each of the lexicons. All positive sentiment scores for “Bing,” “AFINN,” and “NRC” lexicons were recoded to 1 and all negative sentiment scores were recoded to –1. Weekly average scores were calculated to reflect the average sentiment of tweets in a given week. We used the “plot_ly” function in the “plotly” package to visualize changes in weekly sentiment over time.

Supplemental Nonlexicon-Based Approach
ChatGPT (version 3.0) is a next-generation artificial intelligence–based chatbot optimized for using natural language processing to generate responses to user input [32]. We asked ChatGPT (version 3.0) to analyze the overall sentiments of a subset of 150 randomly sampled tweets, with 50 tweets each from our pre-, during-, and postpandemic data sets. We entered, “Can you provide the overall sentiments of the following tweets?” into the query box. ChatGPT responded: “I’d be happy to help you analyze the overall sentiments of the tweets you’ve provided” along with its conclusions on sentiment. We analyzed the sentiment of 50 tweets per data set which was below the maximum size data set allowed for the free version of ChatGPT.

Qualitative Coding
We explored the background literature related to cancer and other health conditions to identify a codebook based on directed content analysis for our project [13,33-36]. We identified Sutton et al [13] for lung cancer messages as the codebook that was most relevant and related to our sample of pediatric cancer tweets. We conducted a directed content analysis [37] using codes and coding definitions from Sutton et al’s codebook. Further, 2 of the authors (NL and AO) coded tweets in sets of 10 to iteratively refine and adapt the Sutton et al codebook and definitions to correspond to pediatric cancer-related tweets. We expanded the preliminary codebook to include the emergence of 7 new coding categories that did not exist in the original codebook. We also added a not enough information coding category for tweets that were ambiguous and could not be coded. Furthermore, 2 of the authors (NL and AO) met weekly to discuss and address codebook discrepancies, and further refine the codebook. The entire authorship team met to discuss codebook development and refinement until it was stable and finalized. The same 2 coders (NL and AO) used the final codebook of content of tweets (adapted from Sutton et al [13]; Textbox 1) to independently double-code all 800 of the randomly sampled tweets.

https://cancer.jmir.org/2024/1/e52061
**Textbox 1.** Final version of qualitative codebook of content of tweets.

<table>
<thead>
<tr>
<th>Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes research on cancer at any point in the continuum, including study results, study in progress, conference presentations, journal publications, research gaps, news publications describing recent findings, and researcher profiles. Any media source, for example, internet blogs, WebMD, or consumer-focused articles apply.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that promotes awareness of cancer (eg, fundraising and prevalence), discusses potential symptoms and signs of cancer, activism, philanthropy, inequities, books, or memoirs about the cancer experience, or makes general references to cancer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Policy and the law</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text about insurance, benefits, legal issues, public policy, and government funding. Code policy and the law only if the tweet does not contain additional content that would lead you to double-code as awareness or another code.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pharmaceuticals and drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that mentions a generic or brand name drug or a pharmaceutical firm.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prevention and risk information</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes cancer risk, behaviors that increase risk (eg, smoking and environmental causes), and behaviors that reduce risk or prevent cancer (eg, healthy diet and smoking cessation).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Early detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes screening tests (eg, low-dose computed tomography), warning signs, early symptoms, and family history.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that contains information about a diagnosis, such as tests (eg, imaging, tests, and biopsy) and results (eg, malignant or benign).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes attempts to medically remove or alter cancer or cancer symptoms (eg, chemotherapy and surgery), discusses treatment of symptoms, references individuals receiving treatment (eg, “fighting cancer”), or information about potential treatments.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Survivorship</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes life after cancer treatment, including remission, and long-term effects of treatment.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mental health</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes the impact of the cancer journey on mental health, mental health treatment or resources, and mental health support.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>End of life</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that discusses cancer-related deaths and legacy. Supportive messages, remembrances, and condolences regarding a patient who died. Parents tweeting about their own children who died of cancer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Personal experiences</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that mentions a personal experience with cancer, including messages about the self and others who have experienced cancer or are worried about cancer. Includes publicized memoirs. If unclear identity of the tweet author (eg, patient, caregiver, and provider), only code personal experiences.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient experiences</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text from individuals with pediatric cancer diagnosis regarding self-experiences. Double-code with personal experiences.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caregiver experiences</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text from caregiver of pediatric cancer regarding personal experiences. Double-code with personal experiences.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health status</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Text that describes current health status, illness progression, and related effects (eg, worries, concerns, and hope).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social support</th>
</tr>
</thead>
</table>
Interrater reliability was calculated as the percent agreement between raters before consensus meetings. Consensus conversations occurred weekly and we referenced the codebook to resolve any discrepancies. The qualitative data were analyzed using DeDoose (Sociocultural Research Consultants, LLC) software for code frequency counts and code co-occurrences. Data visualization of codes was represented by a word cloud generated in DeDoose.

Research Team
Authors' backgrounds included health services research (NL, XZ, AO, HW, and KB), digital health research (NL, XZ, and AO), analytics (XZ), implementation science (NL), clinical psychology (NL, XZ, and AO), pediatric psychology (NL and XZ), bioethics (KB), qualitative research (NL, XZ, and KB), and psychosocial oncology research (NL, AO, HW, and KB).

Results
Sentiment Analyses
Cancer was the most commonly mentioned word, as it was included in all hashtags that were used to collect the tweets. Excluding “cancer,” “Bing” lexicon-based sentiment analyses revealed that there were more positive words than negative words in the extracted tweets (Figure 1). The “Bing” lexicon was based on the largest lexicon among our 3 lexicons and was able to analyze the largest number of tweets. The 5 most commonly observed positive words in the “Bing” lexicon and our data set were “support,” “love,” “amazing,” “heaven,” and “happy.” The 5 most commonly observed negative words in the “Bing” lexicon and our data set were “grief,” “risk,” “hard,” “abuse,” and “miss.” Analyses from the “NRC” lexicon showed that most tweets were categorized under the sentiment types of “positive” (N=138,752), “trust” (N=101,036), and “anticipation” (N=100,635). Figure 2A-C displays weekly sentiment scores over a time period representative of the impact of the COVID-19 pandemic for the “Bing,” “AFINN,” and “NRC” lexicons. The sentiment was overall positive. These findings were consistent across lexicons. Based on responses from ChatGPT (version 3.0), the randomly selected subsamples from the pre-, during-, and postpandemic periods demonstrated overall positive sentiment (Textbox 2). Although ChatGPT (version 3.0) analysis was exploratory, findings were consistent with our lexicon-based analyses.
Figure 1. Most common positive and negative words using the Bing lexicon.
Figure 2. Weekly sentiment using different lexicons from December 11, 2019, to May 7, 2022, a time period representative of the impact of the COVID-19 pandemic. (1) Pre–COVID-19 pandemic, (2) during–COVID-19 pandemic, and (3) post–COVID-19 pandemic.

Textbox 2. Overall sentiments provided by ChatGPT in supplemental analysis.

Pre–COVID-19 pandemic
- The overall sentiments in the provided texts are predominantly positive or neutral. The texts largely revolve around messages of support, encouragement, and raising awareness for childhood cancer, which are inherently positive causes. There is an emphasis on helping and supporting children with cancer and celebrating their milestones. Overall, the texts convey sentiments of empathy and goodwill toward individuals affected by childhood cancer, making the overall sentiment positive and empathetic.

During–COVID-19 pandemic
- Overall, the sentiment in most of these tweets is positive or neutral, as they primarily involve raising awareness, requesting support, or expressing gratitude for contributions to childhood cancer causes.

Post–COVID-19 pandemic
- Overall, this collection of tweets has a predominantly positive and neutral sentiment with some mixed and concerned sentiments. The positivity in these tweets largely stems from support for childhood cancer-related causes and achievements in the field.

Qualitative Coding
Percent agreement between coders was high (91%) before consensus meetings. Consensus meetings resolved all coding discrepancies. The top 5 codes were awareness (N=449), personal experiences (N=166), research (N=60), caregiver experiences (N=54), patient experiences (N=53), policy and the law (N=52), treatment (N=21), end of life (N=21), pharmaceuticals and drugs (N=17), and survivorship (N=15). Data visualization of code applications in a word cloud using DeDoose software is presented in Multimedia Appendix 2.

Twitter users predominantly used the social media platform to promote public awareness of pediatric cancer. In addition, Twitter users frequently use the social media platform to share
personal experiences with cancer. Many accounts were from the firsthand perspectives of caregivers of patients with pediatric cancer in active treatment, bereaved parents, and from patients or survivors of cancer themselves. Twitter was frequently used for health knowledge dissemination of research findings (topics included cancer prevention and risk information, early detection, diagnosis, treatment, and survivorship). Twitter was also frequently used to call attention to and lobby for government programs and federal policies that support pediatric cancer treatment and affordable medical care. Example tweets for the top 10 codes are presented in Textbox 3.

DeDoose software displays the frequency of co-occurring codes using a heat map. Most frequently co-occurring codes are in red, moderately frequent co-occurring codes are in green, and less frequently co-occurring codes are in blue. The code co-occurrence chart of the top 10 codes is presented in Figure 3. Overwhelmingly, awareness and personal experiences were the most frequently cocoded (110 times). Moderately frequent co-occurring codes were caregiver experiences and personal experiences (cocoded 54 times), personal experiences and patient experiences (cocoded 53 times), awareness and patient experiences (cocoded 36 times), and awareness and caregiver experiences (cocoded 34 times).
Textbox 3. Example of paraphrased and deidentified tweets for Top 10 coding categories.

Awareness
- GOLD is the symbolic color for #ChildhoodCancerAwareness. Wearing GOLD to show your support for our children! From head-to-toe, we want to see how gold you can be for Spirit Day! Tag us in your photos.
- We are grateful for the impact on the #pediatriccancer world! Every donation makes a difference.

Personal experiences
- Please help get [this] story out there. #CancerSucks #ChildhoodCancerAwareness. she needs our help!!! RT and donate if you can. Thank you! *Awareness
- FAMily Update» [Child’s name] is back in the hospital. The medical team has ordered a 24 hour [inpatient stay]. #FAM #FightingAllMonsters #ChildhoodCancer *Caregiver Experiences; Health Status

Research
- Brain and spinal cord tumors are the 2nd most common cancers in children. In honor of #ChildhoodCancerAwareness month, here’s a look at recent #Umich discoveries to help treat brain cancer in kids. *Awareness; Treatment
- With more than 10,000 experts worldwide and nearly 100 active clinical trials across the spectrum of childhood cancers, COG is committed to ending childhoodcancer as we know it. #ChildhoodCancerAwareness #ChildhoodCancerAwarenessMonth *Treatment

Caregiver experiences
- I’m involved with many amazing groups for bereaved parents (like myself) and many other #ChildhoodCancer groups. I’d love to see you join us. *End of Life: Personal Experiences
- I am blessed to be the [caregiver] of one of the toughest kids in the world. Love you. #InternationalChildhoodCancerDay #teensvscancer *Personal Experiences

Patient experiences
- [Child’s name] has [medical event] which landed him in the ER. Shout out your loudest prayers and coolest thoughts for [child’s name] so he can return home #FAM #FightingAllMonsters #ChildhoodCancer *Personal Experiences; Social Support
- Agree friends get nervous of saying wrong thing so tend to say nothing I was lucky had a couple of mates there throughout. *Personal Experiences; Social Support
- [Child’s name] finishes his radiotherapy tomorrow!! He put a smile on everyone’s face with his [resilience/playfulness]! *Personal Experiences; Health Status

Policy and the law
- Please do not allow the Creating Hope Reauthorization Act S.4010 to die on YOUR watch! It has produced 28 drugs for rare pediatric diseases, My [child] received 1. *Pharmaceuticals & Drugs; Personal Experiences; Caregiver Experiences
- Please join Congressman Peter Welch VT-0 as a cosponsor of HR 6556 Gabriella Miller Kids First Research Act 2.0. No taxpayer funds required to help #ChildhoodCancer & rare diseases #GMKF2
- We need non-toxic therapies developed for #ChildhoodCancer which is not the same as adults. In the last 10 years, kid’s cancers received only 4% of the budgeted govt research. Will you help? *Awareness

End of life
- This time of year can be especially difficult for those who are grieving, esp. for parents who have lost a child. This is my [personal experience], Here is some advice to help us get through. Please RT. #ChildhoodCancer #Grief #Grieving *Mental Health; Personal Experiences; Caregiver Experiences
- Nothing will ever be so awful as [child’s death]. I am very grateful to the NHS and @TeenageCancer for their efforts but they just couldn’t save him. *Personal Experiences; Caregiver Experiences
- Missing [child’s name] today. Please lend some support to this petition to fund research into childhood cancers #ChildhoodCancer #BrainTumourCharity *Awareness; Personal Experiences

Treatment
- A novel #CARTcelltherapy has shown promising early results in #children with #neuroblastoma, a rare form of ChildhoodCancer. #CancerImmunotherapy *Pharmaceuticals & Drugs
- Clear guidance for stem cell transplant patients, those who have had abdominal radiotherapy, and those who have had total body irradiation as part of transplant #coronavirus #COVID19 #childhoodcancer

Pharmaceuticals and drugs
• A NFCR partner working to advance new therapies for childhood cancer, @OncoHeroesBio, recently announced that the @US_FDA has granted Orphan Drug Designation and Rare Pediatric Disease Designation to one of its drugs. #Together4ACure

• Check out the article: “Leaving no child behind in the fight against cancer” A very good explanation on the current issues we face in ChildhoodCancer drug development, as well as recommendations to solve the current issues! @SIOPEurope

• #ACCELERATEcure (virtual) Annual Conference 2021 - REGISTRATION open next week! Looking forward to welcoming you to discover latest developments worldwide in #ChildhoodCancer drug development! Join us *Research

Survivorship

• The list of potential side effects of #ChildhoodCancer treatments includes future fertility problems, visual loss, dental complications. With the right testing these side effects can be guarded against #ChildhoodCancerAwarenessMonth #Pharmacogenomics *Awareness; Treatment

• Substantial progress has been made against the most common types of pediatric cancers and overall survival rates are up, but more hard work remains so more children with cancer not only survive but thrive. #GoldTogether #ChildhoodCancer *Treatment

Figure 3. Code co-occurrence chart of top 10 coding categories generated by DeDoose.

<table>
<thead>
<tr>
<th></th>
<th>Awareness</th>
<th>Caregiver experiences</th>
<th>End of life</th>
<th>Patient experiences</th>
<th>Personal experiences</th>
<th>Pharmaceuticals and drugs</th>
<th>Policy and the law</th>
<th>Research</th>
<th>Survivorship</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness</td>
<td>0</td>
<td>34</td>
<td>11</td>
<td>36</td>
<td>110</td>
<td>3</td>
<td>17</td>
<td>12</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Caregiver experiences</td>
<td>34</td>
<td>0</td>
<td>12</td>
<td>3</td>
<td>54</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>End of life</td>
<td>11</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Patient experiences</td>
<td>36</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>33</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Personal experiences</td>
<td>110</td>
<td>54</td>
<td>19</td>
<td>53</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Pharmaceuticals and drugs</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Policy and the law</td>
<td>17</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Research</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>15</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Survivorship</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Treatment</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Discussion

Principal Results

The purpose of this study was to examine the communication content of pediatric cancer–related tweets and the public sentiment of pediatric cancer tweets. The sentiment of tweets on pediatric cancer was overall positive, revealing the supportive, hopeful, and inspirational messages relayed by patients, caregivers, and other relevant stakeholders. Our findings were consistent with the only other study to examine the public sentiment of pediatric cancer–related tweets [17].

Despite previous research showing predominantly negative sentiments globally during the COVID-19 pandemic [38], our study describes a positive sentiment of pediatric cancer–related tweets during the COVID-19 pandemic. We found that pediatric cancer–related tweets predominantly focused on raising awareness about pediatric cancer and disseminating health knowledge. We found that both patients with pediatric cancer or survivors and their caregivers frequently used Twitter to provide updates on their health status, for social support, and to share messages of hope.

We only identified 2 studies thus far that have examined the pediatric cancer experience on Twitter. The current study expanded on the growing body of literature on social media use in patients with cancer, and early research on pediatric cancer-related use of Twitter. Previous research studies have discussed the importance of including caregiver experiences in addition to those of the patient for individual and family-based well-being and adaptive coping [39-41]. Our findings were consistent with previous studies that have shown that cancer-related tweets center on health communication and social support [36]. Similar to previous studies, there was a diverse array of Twitter users, representing perspectives from patients, family members, oncology providers, and health care.

https://cancer.jmir.org/2024/1/e52061
pharmaceutical, nonprofit, and other organizations [2,36,42,43]. Additional studies may use combined sentiment analysis and qualitative approaches to better understand pediatric cancer communication and support resources on Twitter and other popular social media platforms. The current and future studies can help inform the development of novel patient- and caregiver-based social media interventions to improve health knowledge, change health behaviors, and improve health outcomes.

Limitations
Our study has several limitations. First, our analysis included only tweets in the English language which limits generalizability to populations that do not speak English as a primary language. Second, we analyzed tweets that contained prespecified keywords (ie, hashtags) and may have missed other pediatric cancer–related tweets during the specified study period. Third, we only examined social media use on Twitter which may differ from usage on other popular social media platforms. Fourth, we were unable to identify the account type (organization vs individual, patient or caregiver vs researcher) and extract sociodemographic information of users; this information may have further informed our research findings and the conclusions drawn. Fifth, our qualitative content analysis of Twitter only included a random sample of tweets from the full data set which may not be representative of all pediatric cancer–related tweets during the specified timeframe of our analysis. Sixth, our lexicon-based approaches have inherent limitations. Despite using multiple sentiment lexicons in our analyses, such approaches analyze sentiment based on individual words. We did not expand contractions and apply stemming in our analyses as they were not available in “saotd,” the statistical package we used for data preprocessing. The lack of expanding contractions and applying stemming may have reduced the number of analyzable words and tweets. Twitter users commonly use contractions, abbreviations, slang, and sarcasm. Thus, we conducted supplemental analyses using ChatGPT (version 3.0), a next-generation artificial intelligence optimized for natural language processing, to validate our findings. Although exploratory, these findings were consistent with lexicon-based approaches. Research should further investigate the use of other recent innovative nonlexicon-based approaches that analyze entire sentences, such as embedding-based approaches or transformer-based approaches to analyze tweets related to pediatric cancer. Seventh, our data were global tweets but our specified “pre-,” “during-,” and “post-” pandemic time periods were based on United States lockdowns and timelines. We acknowledge that pandemic timelines differ within the United States and certainly globally. Nonetheless, we think it is important to include all tweets regardless of geographic location for representativeness of experiences due to the pandemic being a global crisis.

Conclusions
Acute, ongoing, and evolving pediatric medical traumatic stress impacts the child in the context of their family, which emphasizes the importance of incorporating the perspectives and experiences of caregivers and other family members [44,45]. Social media use by patients with pediatric cancer, their families, and their medical providers has been well-described [46]. Uses and benefits include opportunities for social support, building collaborative networks, dissemination of health-related information, and treatment recommendations [46]. Researchers have increasingly turned to sentiment and content analyses of Twitter to capture real-time experiences of patients with a range of health conditions and other relevant stakeholders. Such research has included analyses of tweets about various cancer diagnoses.

Twitter, as a popular social media platform, may serve as an effective means for researchers to examine pediatric cancer communication and public sentiment around the world. This study used both quantitative and qualitative methods to examine the pediatric cancer experience on Twitter. Despite the global mental health crisis during the COVID-19 pandemic, we found overall positive sentiment of pediatric cancer–related tweets over a time period representative of the COVID-19 pandemic. The content of pediatric cancer tweets was posted by a range of users and centered on the delivery of health and treatment information, seeking and providing social support, and the amplification of awareness of pediatric cancer. Twitter may serve as a powerful platform for rapid communication with survivors of pediatric cancer and their caregivers, and facilitate the widespread dissemination of patient- and caregiver-targeted behavioral health interventions to improve well-being and quality of life. This would be well-matched to pediatric cancer survivors’ and their caregivers’ current preferences in social media use.

Acknowledgments
NL was supported by the National Cancer Institute of the National Institutes of Health (K08CA263474). The opinions herein represent those of the authors and not necessarily the funders.

Data Availability
The publicly available data underlying this paper were accessed from Twitter. The derived data generated in this research will be shared on reasonable request to the corresponding author.

Authors’ Contributions
Conceptualization was done by NL, XZ, AO, and KB. Data were curated by NL, XZ, AO, and HW. Formal analysis was handled by NL, XZ, and KB. Investigation was done by NL, XZ, AO, HW, and KB. Methodology was handled by NL, XZ, and KB.
Project administration and supervision were by NL. Visualization was done by XZ and HW. Writing of the original draft was by NL and XZ. Review and editing of the writing was by NL, XZ, AO, HW, and KB.

Conflicts of Interest
XZ has received consulting payments from FirstThen Inc for work unrelated to this paper. The other authors declare no conflict of interest.

Multimedia Appendix 1
An example of the query and timeline information using the hashtag #teenagecancer.

Multimedia Appendix 2
Data visualization of code applications in a word cloud using DeDoose software.

References


**Abbreviations**

DHHS: Department of Health and Human Services
Digital Smoking Cessation Intervention for Cancer Survivors: Analysis of Predictors and Moderators of Engagement and Outcome Alongside a Randomized Controlled Trial

Rosa Andree¹, MSc; Ajla Mujcic², PhD; Wouter den Hollander¹, PhD; Margriet van Laar¹, PhD; Brigitte Boon³,⁴,⁵, PhD; Rutger Engels⁶, PhD; Matthijs Blankers¹,⁷, PhD

¹Trimbos Institute, Netherlands Institute of Mental Health and Addiction, Utrecht, Netherlands
²PsyQ, Parnassia Groep, The Hague, Netherlands
³Siza, Center for Long-term Care for People with Disabilities, Arnhem, Netherlands
⁴Academy Het Dorp, Research & Advisory on Technology in Long-term Care, Arnhem, Netherlands
⁵Tranzo, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, Netherlands
⁶Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, Netherlands
⁷Department of Research, Arkin Mental Health Care, Amsterdam, Netherlands

Corresponding Author:
Rosa Andree, MSc
Trimbos Institute
Netherlands Institute of Mental Health and Addiction
Da Costakade 45
Utrecht, 3521 VS
Netherlands
Phone: 31 30 29 59 267
Email: randree@trimbos.nl

Abstract

Background: Recent studies have shown positive, though small, clinical effects of digital smoking cessation (SC) interventions for cancer survivors. However, research on associations among participant characteristics, intervention engagement, and outcomes is limited.

Objective: This study aimed to explore the predictors and moderators of engagement and outcome of MyCourse-Quit Smoking (in Dutch: “MijnKoers-Stoppen met Roken”), a digital minimally guided intervention for cancer survivors.

Methods: A secondary analysis of data from the randomized controlled trial was performed. The number of cigarettes smoked in the past 7 days at 6-month follow-up was the primary outcome measure. We analyzed interactions among participant characteristics (11 variables), intervention engagement (3 variables), and outcome using robust linear (mixed) modeling.

Results: In total, 165 participants were included in this study. Female participants accessed the intervention less often than male participants (B=–11.12; \( P=.004 \)). A higher Alcohol Use Disorders Identification Test score at baseline was associated with a significantly higher number of logins (B=1.10; \( P<.001 \)) and diary registrations (B=1.29; \( P<.001 \)). A higher Fagerström Test for Nicotine Dependence score at baseline in the intervention group was associated with a significantly larger reduction in tobacco use after 6 months (B=–9.86; \( P=.002 \)). No other associations and no moderating effects were found.

Conclusions: Overall, a limited number of associations was found between participant characteristics, engagement, and outcome, except for gender, problematic alcohol use, and nicotine dependence. Future studies are needed to shed light on how this knowledge can be used to improve the effects of digital SC programs for cancer survivors.

Trial Registration: Netherlands Trial register NTR6011/NL5434; https://onderzoekmetmensen.nl/nl/trial/22832

(JMIR Cancer 2024;10:e46303) doi:10.2196/46303

KEYWORDS
smoking cessation; cancer survivors; engagement; digital intervention; eHealth; smoking; intervention; randomized controlled trial; predictor; RCT; smoking; smoker; addict; cessation; quit; cancer; oncology

https://cancer.jmir.org/2024/1/e46303
Introduction

Background

In the past decade, digital interventions have commonly been used to target addictive behaviors, including smoking cessation (SC). Several systematic reviews have shown that these SC interventions can be effective, albeit with generally small effect sizes [1-4]. For example, the Cochrane review by Taylor et al [4] showed that the use of web-based SC interventions resulted in significantly higher rates of smoking abstinence compared to nonactive control groups, 6 months after randomization (risk ratio=1.15). Cancer survivors are a growing population who can benefit considerably from SC. Yet, the prevalence of people who smoke is about the same as in the general population, and research on effective digital SC interventions for cancer survivors is scarce [3]. Accordingly, not much is known about active ingredients or engagement factors of SC interventions targeting cancer survivors [5], despite engagement being an important moderator of the effect of digital SC interventions [6]. It is therefore important to look more closely into the predictors and moderators of engagement and outcome among this target group.

Although the primary effects of digital SC interventions are moderately positive on average, there is room for improvement. One possible explanation for the modest effects of digital SC interventions is the generally low adherence rates. Taylor et al [4] found that 18 out of 34 web-based SC studies had more than 50% attrition at follow-up. Analyzing whether the uptake of specific intervention components is related to better intervention outcomes increases the understanding of primary intervention effects [7]. Some studies on addictive behaviors investigated the relationship between intervention engagement and outcome [8-10]. A study by Perski et al [8] found that participants who completed more (varied) exercises had 64% higher odds of SC compared to participants who almost exclusively set an SC goal. Siemer et al [10], examining adherence to a blended SC intervention, revealed a dose-response relationship between the number of executed activities and smoking abstinence. Another study by Ramos et al [9] also found that intervention engagement, in terms of number of logins, forum visits, and number of participation badges, was a strong predictor of successful SC. Not all studies have shown that intervention engagement predicts intervention effectiveness, even contradictory effects are found. For example, Smith et al [11] showed that engagement with particular components of a digital SC intervention can be counterproductive when the content does not fit the participants’ needs.

Behavior change programs targeting SC notoriously encounter challenges when trying to reach target groups with the highest smoking rates (eg, groups with lower socioeconomic status [12] and groups with low literacy [13]). In addition, it could be useful for the improvement of intervention content and implementation to identify which characteristics predict engagement. This will help to improve the content and design of the intervention for the right target group [7,14]. Several studies have related digital SC intervention use to participant characteristics [8,15-17]. These studies showed that digital SC interventions were used longer or more frequently by older participants [8,15] and women [16]. Participants who had lower education, smoked more heavily, and had depressive symptoms were found to be less engaged with the digital SC intervention [17].

There is some evidence on the effects of digital SC interventions for cancer survivors. For example, a meta-analytic study by Mujic et al [3] showed that digital and nondigital distance-based SC interventions for cancer survivors led to significantly reduced smoking rates compared to baseline (risk difference=0.29). However, research on the predictors and moderators of engagement and outcome of digital SC studies for cancer survivors is limited, while cancer survivors are a growing and diverse population [18]. A pilot study by Bricker et al [18] of an application on SC for cancer survivors showed greater acceptability, use, and effectiveness when compared to the national SC app for the general population.

This Study

In this study, we aim to investigate the predictors and moderators of engagement and outcome of a minimally guided digital intervention for cancer survivors called MyCourse-Quit Smoking (in Dutch “MijnKoers-Stoppen met Roken”) in a secondary analysis. The main effects study did not find a differential effect on SC between intervention and control at 6 months. In both groups, around a quarter abstained from smoking, and the number of cigarettes smoked was cut back by half [19]. With this secondary analysis, we aim to answer the following research questions: (1) Are participant characteristics related to intervention engagement at 6-month follow-up? (2) Is intervention engagement associated with tobacco use at 6-month follow-up? (3) Are participant characteristics related to tobacco use at 6-month follow-up?

Methods

Design

For this paper, an exploratory secondary data analysis was carried out using data from a randomized controlled trial on the MyCourse-Quit Smoking digital intervention. The data used for this study were collected between November 2016 and September 2019.

Ethical Considerations

Ethics approval for the trial was acquired from an accredited medical research and ethics committee in The Netherlands (Toetsingscommissie Wetenschappelijk Onderzoek Rotterdam e.o. NL55921.101.16). Participants provided digital informed consent before inclusion in the trial [20]. Data were deidentified before processing or analysis. Identifying data were stored separately from research data. For each completed follow-up assessment, they were reimbursed €25 (approximately US $30).

Procedure

A web-based screening questionnaire on the study website determined whether people were able to participate in the study. Eligible participants received an informed consent form via mail and had 30 days to sign the form. In the meantime, participants had the possibility to contact the research team or an independent physician for more information. After signing...
the informed consent form, they were asked to fill out a baseline questionnaire and were allocated to either the MyCourse group or the control group. Individuals in the control group were provided access to a noninteractive web-based informational brochure regarding the hazards associated with smoking and strategies for SC. The informational content encompassed both general SC information and content tailored to the unique needs of cancer survivors. Follow-up measurements were conducted at 3, 6, and 12 months after randomization. The study was conducted completely over the web, but after continued nonresponse, participants received a reminder by telephone. A more extensive description of the randomized controlled trial study procedures can be found elsewhere [20].

**Participants**

For the study, 165 adults who were diagnosed with any form of cancer in the past 10 years were recruited. Other eligibility criteria included having a PC or laptop in addition to an internet connection at home, having smoked 5 or more cigarettes per day in the past 7 days, having the intention and ability to participate in the 12-month study, and having the intention to quit smoking cigarettes. People who were pregnant; had insufficient mastery of the Dutch language; or self-reported suicidal ideation, dementia, severe depression, severe alcohol dependence, or acute psychosis were not eligible to participate in the study.

**MyCourse Intervention**

MyCourse-Quit Smoking is a digital minimally guided intervention that provides support for SC among cancer survivors. The intervention is based on empirically evaluated therapeutic approaches for SC in the general population: cognitive behavioral therapy [21], motivational interviewing [22], and acceptance and commitment therapy [23]. The intervention can be accessed via PC, tablet, and smartphone. At first login, participants receive instructions to set up a quit plan and gain access to 13 different exercises, information about smoking, quitting, and cancer, a web-based diary to track their tobacco use, and a peer support platform [20]. Exercises focused on different topics including previous experiences, high-risk situations, self-control measures, reinforcement, relapse prevention, and acceptance and commitment therapy. For the complete structure of the intervention, see Figure 3 in the protocol paper [20]. After the first login, all parts of the intervention could be accessed, and participants were free to choose how often and which parts of the intervention they wanted to use. Participants were only advised to use the intervention daily for 4 weeks.

**Measures**

The primary outcome measure in this study was the number of cigarettes smoked in the past 7 days at 6-month follow-up. Intervention engagement was measured using 3 indicators: the number of logins into the MyCourse-Quit Smoking intervention, the number of self-monitoring registrations of smoking urges and smoked cigarettes, and the number of completed intervention exercises. The following participant characteristics were extracted from the participant records: gender, age, educational level (higher or lower, where the minimum for the higher educational level was an academic university or university of applied sciences degree), and living situation (alone or together). We specifically looked at the presence of lung cancer and breast cancer (yes or no) among the participants in the analyses because lung cancer has a direct relationship with smoking and breast cancer was the most common type of cancer in the sample. In addition, patients with cancer at other sites were included in the analyses. Furthermore, the number of cancer sites (1 or >1) distinguished participants who reported that they had received multiple cancer diagnoses. The severity of nicotine dependence was measured by the 6-item Fagerström Test for Nicotine Dependence (FTND) [24]. Problematic alcohol use was measured using the Alcohol Use Disorders Identification Test (AUDIT) [25], a 10-item questionnaire on alcohol consumption patterns and problems experienced due to alcohol consumption. The AUDIT score was included as a variable because research has shown that people with a high risk of problematic alcohol use have a harder time quitting smoking and may benefit from different types of SC treatment [26,27]. The EQ-5D was used to measure the quality of life [28]. Comorbid anxiety, depression, and somatic symptoms were indicated using the Brief Symptom Inventory-18 questionnaire [29].

**Statistical Analysis**

**Imputation of Missing Data**

Missing data for primary (ie, cigarettes smoked in the past 7 days) and secondary (ie, participant characteristics) outcome measures were multiple imputed (number of imputations=50) based on the intention-to-treat principle using the predictive mean matching method from the mice package in R (R Foundation for Statistical Computing) [29]. At the 6-month follow-up, the nonresponse rate (ie, participants who did not complete the 6-month questionnaire) was 27.7% (23/83) in the intervention group and 25.6% (21/82) in the control group. Data on intervention usage were not imputed. For the analyses containing engagement measures, participants who did not log in once were excluded.

**Regression Analyses**

Data were analyzed using R [30]. Bonferroni correction was applied in all analyses. The association between intervention engagement and participant characteristics within the intervention group was analyzed with a robust linear regression using the MASS package [31]. Whether participant characteristics and intervention engagement predicted intervention outcome within the intervention group was analyzed using robust linear mixed modeling (RLMM) with a random intercept using the robustlmm package [32]. RLMM is an effective analytical approach to account for outliers or skewed data [32]. The moderation analyses to investigate whether the study condition moderated the effect between participant characteristics and outcome were performed using RLMM with a random intercept and study condition × participant characteristics as the interaction term. This analysis is performed to assess whether the study condition (ie, being in the intervention group compared to the control group) moderates the association between participant characteristics and outcome. Model estimates, 95% CIs, and P values are reported. All base
case analyses followed the intention-to-treat principle and used multiple imputed data sets. Sensitivity analyses were performed using observed data only.

**Results**

**Sample Characteristics**

Participant characteristics are shown in Table 1. The majority of participants were female (136/165, 82.4%), the mean age was 54.2 (SD 11.2) years, 29.1% (48/165) were living alone, and 41.2% (68/165) had completed higher education. On average, participants had smoked for 34.5 (SD 12.0) years and smoked 100 (SD 51.2) cigarettes per week. The main clinical effects of the MyCourse intervention and the results of the cost-effectiveness analysis can be found elsewhere [19].

**Table 1.** Baseline characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>MyCourse (n=83)a</th>
<th>Control (n=82)</th>
<th>Total (N=165)a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>70 (84.3)</td>
<td>66 (80.5)</td>
<td>136 (82.4)</td>
</tr>
<tr>
<td>Male</td>
<td>13 (15.7)</td>
<td>16 (19.5)</td>
<td>29 (17.6)</td>
</tr>
<tr>
<td><strong>Age (years), mean (SD)</strong></td>
<td>55.0 (12.1)</td>
<td>53.3 (10.3)</td>
<td>54.2 (11.2)</td>
</tr>
<tr>
<td><strong>Higher education, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25 (30.1)</td>
<td>19 (23.2)</td>
<td>44 (26.7)</td>
</tr>
<tr>
<td>No</td>
<td>49 (59.0)</td>
<td>48 (58.5)</td>
<td>97 (58.8)</td>
</tr>
<tr>
<td><strong>Living situation, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td>22 (26.5)</td>
<td>26 (31.7)</td>
<td>48 (29.1)</td>
</tr>
<tr>
<td>Living together</td>
<td>61 (73.5)</td>
<td>56 (68.3)</td>
<td>117 (70.9)</td>
</tr>
<tr>
<td><strong>Smoking behavior, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years smoked</td>
<td>34.4 (11.8)</td>
<td>34.6 (12.2)</td>
<td>34.5 (12)</td>
</tr>
<tr>
<td>Number of cigarettes in past 7 days</td>
<td>101.8 (54.3)</td>
<td>98.2 (48.2)</td>
<td>100 (51.2)</td>
</tr>
<tr>
<td><strong>FTND</strong>b</td>
<td>4.9 (2.4)</td>
<td>4.9 (2.3)</td>
<td>4.9 (2.4)</td>
</tr>
<tr>
<td><strong>Drinking behavior</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drank alcohol in last month, n (%)</td>
<td>55 (66.3)</td>
<td>55 (67.1)</td>
<td>110 (66.7)</td>
</tr>
<tr>
<td>Number of drinks in past 7 days, mean (SD)</td>
<td>6.9 (13.1)</td>
<td>5.6 (8.7)</td>
<td>6.2 (11.2)</td>
</tr>
<tr>
<td><strong>AUDITc, mean score (SD)</strong></td>
<td>3.7 (5.1)</td>
<td>3.6 (4.2)</td>
<td>3.6 (4.7)</td>
</tr>
<tr>
<td><strong>Cancer diagnosis, d n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>42 (42.9)</td>
<td>33 (34)</td>
<td>75 (38.5)</td>
</tr>
<tr>
<td>Lung</td>
<td>14 (14.3)</td>
<td>9 (9.3)</td>
<td>23 (11.8)</td>
</tr>
<tr>
<td>Uterine</td>
<td>7 (7.1)</td>
<td>12 (12.4)</td>
<td>19 (9.7)</td>
</tr>
<tr>
<td>Head and neck</td>
<td>10 (10.2)</td>
<td>8 (8.2)</td>
<td>18 (9.2)</td>
</tr>
<tr>
<td>Colon</td>
<td>5 (5.1)</td>
<td>5 (5.2)</td>
<td>10 (5.1)</td>
</tr>
<tr>
<td>Other (including bladder, lymphatic, melanoma, skin, kidney, prostate)</td>
<td>20 (20.4)</td>
<td>30 (30.9)</td>
<td>50 (25.6)</td>
</tr>
<tr>
<td><strong>Cancer sites, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>69 (83.1)</td>
<td>68 (82.9)</td>
<td>137 (83)</td>
</tr>
<tr>
<td>2 or 3</td>
<td>14 (16.9)</td>
<td>14 (17.1)</td>
<td>28 (17)</td>
</tr>
</tbody>
</table>

**Participant Characteristics and Intervention Engagement**

Of all 83 participants of the intervention group, 56 people logged into the MyCourse portal at least once and thus were included in the analysis. When comparing the 56 people who logged in at least once with the 27 people who did not log in once at all baseline characteristics mentioned in Table 1, only the number of patients with uterine cancer differed significantly between the 2 groups (P<.05), with 5 patients with uterine cancer who did not log in once and 2 patients with uterine cancer that logged in at least once. In total, 82 participants in the control group were not included in the analysis. Among the 56 MyCourse users, the average number of logins was 21 (SD 41.0; median...
5.5, IQR 3-18.5), the average amount of self-monitoring registrations was 31 (SD 53.9; median 5, IQR 2-22), and the average amount of completed exercises was 6.5 (SD 5.1; median 4, IQR 2-12). As shown in Table 2, female participants showed a significantly lower number of logins in the MyCourse-Quit Smoking intervention than male participants ($P=.004$). The relationship between sex and other indicators of intervention engagement was nonsignificant. Furthermore, a higher AUDIT score at baseline was associated with a significantly higher number of logins ($P<.001$) and diary registrations ($P<.001$) but not with the number of completed exercises ($P=.05$).

### Table 2. The association between baseline participant characteristics and intervention engagement (N=56).

<table>
<thead>
<tr>
<th></th>
<th>Logins</th>
<th>Diary entries</th>
<th>Exercises</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B (95% CI)</td>
<td>P values $^a$</td>
<td>B (95% CI)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.25 (–0.10 to 0.60)</td>
<td>.16</td>
<td>0.40 (–0.16 to 0.96)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (n=48)</td>
<td>–11.12 (–18.70 to –3.55)</td>
<td>$.004 ^b</td>
<td>–11.85 (–24.25 to 0.54)</td>
</tr>
<tr>
<td>Male (n=8)</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Higher education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=40)</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Yes (n=16)</td>
<td>3.08 (–4.10 to 10.26)</td>
<td>.40</td>
<td>4.43 (–4.71 to 13.57)</td>
</tr>
<tr>
<td>Living situation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alone (n=11)</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Together (n=45)</td>
<td>0.32 (–7.46 to 8.11)</td>
<td>.94</td>
<td>0.72 (–10.64 to 12.08)</td>
</tr>
<tr>
<td>FTND$^c$</td>
<td>–0.15 (–1.51 to 1.22)</td>
<td>.83</td>
<td>–0.33 (–2.32 to 1.66)</td>
</tr>
<tr>
<td>EQ-5D$^d$</td>
<td>–4.88 (–21.59 to 11.84)</td>
<td>.57</td>
<td>–9.43 (–33.30 to 14.44)</td>
</tr>
<tr>
<td>BSI-18$^d$</td>
<td>–3.57 (–8.36 to 1.23)</td>
<td>.15</td>
<td>–5.07 (–11.46 to 1.32)</td>
</tr>
<tr>
<td>AUDIT$^e$</td>
<td>1.10 (0.60 to 1.61)</td>
<td>&lt;.001 ^f</td>
<td>1.29 (0.62 to 1.95)</td>
</tr>
<tr>
<td>Diagnosis of lung cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=47)</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Yes (n=9)</td>
<td>–2.19 (–10.48 to 6.09)</td>
<td>.60</td>
<td>–3.01 (–15.04 to 9.02)</td>
</tr>
<tr>
<td>Diagnosis of breast cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=25)</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Yes (n=31)</td>
<td>3.70 (–2.59 to 9.99)</td>
<td>.25</td>
<td>5.94 (–1.87 to 13.74)</td>
</tr>
<tr>
<td>Cancer sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (n=47)</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>2 or 3 (n=9)</td>
<td>–2.75 (–11.29 to 5.79)</td>
<td>.53</td>
<td>–3.88 (–17.73 to 9.98)</td>
</tr>
</tbody>
</table>

$^a$A Bonferroni correction was applied based on 11 tests resulting in an $\alpha$ of .0045.

$^b$Female participants showed a significantly lower number of logins in the MyCourse-Quit Smoking intervention than male participants.

$^c$FTND: Fagerström Test for Nicotine Dependence.

$^d$BSI-18: Brief Symptom Inventory-18.

$^e$AUDIT: Alcohol Use Disorders Identification Test.

$^f$A higher AUDIT score at baseline was associated with a significantly higher number of logins and diary registrations but not with the number of completed exercises.

### Intervention Engagement, Participant Characteristics, and Smoking Behavior

Table 3 shows the outcomes of the analysis on the association between intervention engagement and smoking behavior among the 56 participants who logged in to the MyCourse portal at least once. No significant effects were found between intervention engagement and smoking behavior. Table 3 also shows the association between several participant characteristics and smoking behavior among the 83 participants of the intervention group. The results show that a higher FTND score at baseline is associated with a significantly greater reduction of the 7-day sum of smoked cigarettes after 6 months in the intervention group ($P=.002$). None of the other participant characteristics or measures of engagement predicted smoking behavior at 6 months.
### Table 3. The relationship between participant characteristics and intervention engagement with smoking behavior.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Effect on 7-day tobacco use at 6-month follow-up</th>
<th>P values&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years) (n=83)</strong></td>
<td>B (95% CI)</td>
<td>.39</td>
</tr>
<tr>
<td></td>
<td>0.70 (–0.89 to 2.29)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n=13)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Female (n=70)</td>
<td>–5.71 (–51.07 to 39.65)</td>
<td>.81</td>
</tr>
<tr>
<td><strong>Higher education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=58)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Yes (n=25)</td>
<td>–2.16 (–37.12 to 32.80)</td>
<td>.90</td>
</tr>
<tr>
<td><strong>Living situation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alone (n=22)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Together (n=61)</td>
<td>–1.04 (–38.13 to 36.05)</td>
<td>.96</td>
</tr>
<tr>
<td>FTND&lt;sup&gt;b&lt;/sup&gt; (n=83)</td>
<td>–9.86 (–15.95 to –3.76)</td>
<td>.002&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>EQ-5D (n=83)</td>
<td>17.28 (–64.19 to 98.74)</td>
<td>.68</td>
</tr>
<tr>
<td>BSI-18&lt;sup&gt;d&lt;/sup&gt; (n=83)</td>
<td>1.25 (–27.43 to 29.94)</td>
<td>.93</td>
</tr>
<tr>
<td>AUDIT&lt;sup&gt;e&lt;/sup&gt; (n=83)</td>
<td>2.38 (–0.76 to 5.53)</td>
<td>.14</td>
</tr>
<tr>
<td><strong>Diagnosis of lung cancer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=69)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Yes (n=14)</td>
<td>19.84 (–10.95 to 50.62)</td>
<td>.21</td>
</tr>
<tr>
<td><strong>Diagnosis of breast cancer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=41)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Yes (n=42)</td>
<td>11.77 (–32.27 to 55.81)</td>
<td>.60</td>
</tr>
<tr>
<td><strong>Cancer sites</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (n=69)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>2 or 3 (n=14)</td>
<td>53.77 (13.70 to 93.83)</td>
<td>.009</td>
</tr>
<tr>
<td><strong>Number of logins (n=56)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of logins</td>
<td>–0.13 (–0.55 to 0.30)</td>
<td>.56</td>
</tr>
<tr>
<td><strong>Number of diary entries (n=56)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of diary entries</td>
<td>–0.02 (–0.35 to 0.30)</td>
<td>.88</td>
</tr>
<tr>
<td><strong>Number of exercises (n=56)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of exercises</td>
<td>0.19 (–3.31 to 3.69)</td>
<td>.92</td>
</tr>
</tbody>
</table>

---

<sup>a</sup>A Bonferroni correction was applied based on 14 tests resulting in an α of .004.

<sup>b</sup>FTND: Fagerström Test for Nicotine Dependence.

<sup>c</sup>A higher FTND score at baseline is associated with a significantly greater reduction of the 7-day sum of smoked cigarettes after 6 months in the intervention group.

<sup>d</sup>BSI-18: Brief Symptom Inventory-18.

<sup>e</sup>AUDIT: Alcohol Use Disorders Identification Test.

### Moderation Analysis

Table 4 reports the outcomes of the moderation analysis on the interaction effect of participant characteristics and study condition on the number of cigarettes smoked in the past 7 days among the 165 participants at 6-month follow-up. No significant effects were found in this analysis.
Table 4. Moderation analysis of study condition on the relationship between participant characteristics and smoking behavior.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participant characteristic × randomized controlled trial condition 7-day tobacco use at 6-month follow-up (N=165)</th>
<th>B (95% CI)</th>
<th>P values&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td>1.229 (−1.06 to 3.52)</td>
<td>.29</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n=29)</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (n=136)</td>
<td>−11.7 (−73.76 to 50.36)</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>Higher education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=121)</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (n=44)</td>
<td>−25.21 (−77.18 to 26.76)</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td>Living situation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alone (n=48)</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Together (n=117)</td>
<td>7.989 (−43.59 to 59.57)</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>FTND&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−3.624 (−12.20 to 4.95)</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>EQ-5D</td>
<td>45.117 (−75.26 to 165.49)</td>
<td>.46</td>
<td></td>
</tr>
<tr>
<td>BSI-18&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−12.024 (−55.66 to 31.61)</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>AUDIT&lt;sup&gt;d&lt;/sup&gt;</td>
<td>4.155 (−1.04 to 9.35)</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>Diagnosis of lung cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=142)</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (n=23)</td>
<td>42.98 (−24.59 to 110.55)</td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td>Diagnosis of breast cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (n=90)</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (n=75)</td>
<td>9.25 (−36.89 to 55.39)</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>Cancer sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (n=137)</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 or 3 (n=28)</td>
<td>81.71 (22.50 to 140.91)</td>
<td>.007</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>A Bonferroni correction was applied based on 11 tests resulting in an α of .0045.

<sup>b</sup>FTND: Fagerström Test for Nicotine Dependence.

<sup>c</sup>BSI-18: Brief Symptom Inventory-18.

<sup>d</sup>AUDIT: Alcohol Use Disorders Identification Test.

Discussion

Principal Findings

In this study, we evaluated hypothesized predictors and moderators of intervention engagement and smoking behavior in MyCourse-Quit Smoking, a digital SC intervention for cancer survivors. With regard to the relationship between participant characteristics and intervention engagement, it was found that female participants logged on significantly less often than male participants. This effect should nevertheless be interpreted with caution since the number of male participants in the sample was low (n=8). Moreover, previous research shows that female participants are generally more engaged in digital SC interventions than male participants [16,33-37]. A significant positive association between the baseline AUDIT score and intervention engagement was found; a higher AUDIT score at...
baseline was related to a higher number of logins and diary registrations in the MyCourse intervention. There was no effect of the baseline AUDIT score on the number of completed exercises. Previous studies showed that participants with a higher risk of alcohol dependence had a harder time to quit smoking, and therefore needed more support from the intervention, as demonstrated in several previous studies [26,27]. Toll et al [27] showed that people who drink more heavily were less likely to quit smoking, but problematic alcohol use was not measured. Sells et al [26] pointed out that people with a high risk of problematic alcohol use may need more intensive intervention in order to quit smoking, whereas people with a high risk of problematic alcohol use were defined with an AUDIT score higher than 7. However, in this study, we did not find an effect of the AUDIT score on smoking behavior. Furthermore, participants of the MyCourse-Quit Smoking trial had generally low AUDIT scores (mean 3.6, SD 4.7), and few participants with a score higher than 7 (21/165).

Regarding the association between participant characteristics and smoking behavior, we found that participants of the MyCourse intervention who had higher nicotine dependence scores at baseline showed a greater reduction in the number of smoked cigarettes in the past 7 days at the 6-month follow-up. This negative association between nicotine dependence at baseline and tobacco use at follow-up is a reasonable finding because it is likely that less addicted participants at baseline already smoke fewer cigarettes than highly addicted participants, and therefore, a smaller reduction of cigarettes at 6 months is possible. This finding does not indicate whether heavier nicotine dependence predicts SC, as participants can greatly reduce the number of smoked cigarettes but not enough to completely quit smoking. Previous research shows that, in general, less severe nicotine dependence is associated with a higher SC rate [38,39].

The analyses on the association between intervention engagement (ie, the number of logins, self-monitoring registrations, and exercises) and the outcome did not yield any significant effects. This study showed the overall prevailing pattern of the majority of participants quitting the use of the intervention in the first few days and a smaller group that uses the intervention for a longer period [40]. However, other studies on digital SC interventions have shown a dose-response relationship between intervention engagement and outcome [9,10,41], with higher engagement predicting greater SC rates, although this is sometimes limited to certain engagement measures [34] or with low quality of the evidence due to low follow-up rates [9]. For example, Heminger et al [34] did not find a significant association between program dose and SC, but the use of specific intervention elements (eg, making a pledge toward a smoke-free life and tracking saved money and health benefits gained after quitting) was associated with SC. For future research, it is therefore important to properly define engagement, differentiate between indicators of engagement, and use empirically effective intervention techniques in order to enhance engagement [6].

The moderation analysis did not yield any significant effects. This indicates that being in the intervention group, compared to the control group, does not amplify the effect between any of the participant’s characteristics and tobacco use, and hence no specific participant characteristic renders participants more or less likely to be successful when participating in the MyCourse intervention.

Limitations

The initial study was 80% powered to detect a relative risk of 2.1 in SC [20], while this explored different outcome variables, potential moderator effects herein, and made comparisons other than between treatment arms. Hence, the initial sample size calculation might not be applicable. Post hoc power analyses were omitted, as these would merely reflect the already obtained P value [42]. While the applied Bonferroni correction accounted for multiple comparisons, it might be overly strict in our case [43]. Furthermore, the tendency to overfit data might also be a problem for linear mixed modeling analyses. The study had missing data, which might have caused bias in the results. On the other hand, as a strength of this study, multiple imputation was applied to compensate for the missing values, and the sensitivity analysis did not reveal any substantial differences in the analyses without imputation. Another limitation is the sample size of the analyses for the first research question, especially for the subgroup analyses of sex and living situation. Since some of the categories of these variables had small group sizes, the outcomes of the analyses should be interpreted with caution.

Clinical Implications

The MyCourse intervention is presumably more engaging for people who smoke and people with moderate to high alcohol dependence. Furthermore, this study did not identify any specific subgroups where the MyCourse-Quit Smoking intervention might be particularly effective or ineffective.

Conclusions

This study aimed to provide more insight into predictors and moderators of engagement and outcome for a digital SC intervention targeting cancer survivors. Overall, a limited number of associations was found between participant characteristics, engagement, and smoking behavior. Female participants accessed the intervention less often than male participants, and participants with higher AUDIT scores accessed the intervention more often and had more diary registrations than participants with lower AUDIT scores. Greater nicotine dependence at baseline was associated with a greater reduction in number of cigarettes at 6 months. Future studies in a larger sample and with a preregistered analysis plan are needed to corroborate these findings and shed light on how this knowledge can be used to improve the effects of digital SC programs.
Acknowledgments
This study was funded by the Dutch Cancer Institute (KWF kankerbestrijding; grant TBOS2014–7169). The sponsors did not have influence on the design, data collection, analysis, and interpretation of the data, nor in writing the paper or the decision to submit it for publication. The authors would like to thank Dieder Kesler for her assistance in the recruitment and follow-up assessments of the participants and Yvonne Borghans for her assistance in the follow-up assessments of the participants.

Authors’ Contributions
AM, MvL, BB, RE, and MB contributed to the conception and data collection of the original research. RA, AM, WdH, and MB conceived the research questions and design for this study. ML, BB, and RE commented on or rewrote the design and research questions. RA and WdH performed the statistical analyses. RA wrote the first draft of the manuscript. AM, WdH, MvL, BB, RE, and MB commented on the draft and rewrote sections of the draft. All authors approved the final version of the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Sensitivity analysis of all performed analyses.
[DOCX File, 38 KB - cancer_v10i1e46303_app1.docx ]

References


Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
FTND: Fagerström Test for Nicotine Dependence
RLMM: robust linear mixed modeling
SC: smoking cessation
Assessing the Quality, Privacy, and Security of Breast Cancer Apps for Arabic Speakers: Systematic Search and Review of Smartphone Apps

Dari Alhuwail1,2, PhD; Aisha Alhouti1, MSc; Latifah Alsarhan1, BS

1Information Science Department, College of Life Sciences, Kuwait University, Sabah AlSalem University City, Kuwait
2Health Informatics Unit, Dasman Diabetes Institute, Dasman, Kuwait

Corresponding Author:
Dari Alhuwail, PhD
Information Science Department
College of Life Sciences
Kuwait University
P.O. Box 5969
Sabah AlSalem University City, Safat 13060
Kuwait
Phone: 965 24633214
Email: dari.alhuwail@ku.edu.kw

Abstract

Background: Breast cancer is a widespread disease, and its incidence is rapidly increasing in the Middle East and North Africa region. With the increasing availability of smartphone apps for various health purposes, breast cancer apps have emerged as tools for raising awareness, providing support, and empowering women affected by this disease. These apps offer many features, including information on breast cancer risk factors, self-examination guides, appointment reminders, and community support groups or hotlines. Using apps raises the risk of privacy and security issues, and we hope that examining these features of the apps will contribute to the understanding of how technology can be used to improve these apps and provide insights for future development and improvement of breast cancer apps.

Objective: This study aims to critically review the quality, privacy, and security of breast cancer apps available to Arabic speakers.

Methods: Similar to several recent studies, we used a systematic search for apps available in Google Play and Apple App stores using both the web interface and the built-in native stores installed on smartphones. The search was conducted in mid-December 2022 in Arabic using the following keywords: “breast cancer, breast cancer treatment, breast cancer disease, breast cancer symptoms, breast cancer screening, and breast test”. These preidentified search terms are based on earlier work concerning the top searched breast cancer topics by Arabic speakers through Google’s search engine. We excluded apps that did not have an Arabic interface, were developed for non-Arabic speakers, were paid, needed a subscription, or were directed toward health care workers. The Mobile App Rating Scale was used to evaluate the quality of the apps concerning their engagement, functionality, aesthetics, and information. A risk score was calculated for the apps to determine their security risk factors.

Results: Only 9 apps were included, with most (6/9, 67%) being supported by advertisements and categorized as informational. Overall, the apps had low numbers of downloads (>10 to >1000). The majority of the included apps (8/9, 89%) requested dangerous access permissions, including access to storage, media files, and the camera. The average security score of the included apps was 3.22, while only 2 apps provided information about data security and privacy. The included apps achieved an overall average quality score of 3.27, with individual dimension scores of 4.75 for functionality, 3.04 for information, 3.00 for aesthetics, and 2.32 for engagement.

Conclusions: The limited availability of breast cancer apps available to Arabic speakers should be a call to action and prompt health care organizations and developers to join forces and collaboratively develop information-rich, usable, functional, engaging, and secure apps.

(JMIR Cancer 2024;10:e48428) doi:10.2196/48428
KEYWORDS
apps; Arabic; awareness; breast cancer; consumer health informatics; education; mHealth; mobile health; privacy; quality; security; smartphone; women

Introduction
Female breast cancer is among the most commonly diagnosed cancers worldwide, with a rate of approximately 2.3 million new cases [1]. Over the past decade, breast cancer incidence has been on the rise in the Middle East and North Africa (MENA) region [2]. Breast cancer is the most frequently diagnosed cancer (17.7%-19% of all types of cancer) in the region [3], and it accounts for 30% of all cancer cases [4]. The lack of cancer education and barriers to cancer screening are seen as major health problems [5]. Education is one of the most effective tools in the fight against female breast cancer; it can have positive effects on women’s practices, attitudes, and knowledge of the disease [6]. However, recent studies suggest a lack of resources and poor awareness of breast cancer in women in the MENA region [7]. While the digital delivery of health education content has been on the rise through different digital media (eg, websites and social media), the quality of Arabic content for female breast cancer remains poor [8-10].

Smartphone proliferation, ubiquity, and affordability, as well as the increasing availability of mobile apps, may be the long-awaited for “digiceuticals” or digital therapeutics [11-13]. Today, the number of health apps in smartphone app stores exceeds 325,000 [14] and will continue to rise, with estimates of more than 200 apps being added daily to app stores [15], covering a wide spectrum of health purposes, such as well-being, education, and disease management, including chronic conditions [16]. Several apps are available that can help individuals with breast cancer manage their condition. These apps have been used for the purposes of education [17-19], care management [20,21], prevention [22-24], and well-being [25,26]. These apps can be a valuable resource for individuals with breast cancer, helping them stay informed and connected to their care team and manage the symptoms and challenges associated with the disease.

The use of mobile health (mHealth) apps contributes to improving health literacy and facilitating communication between patients and their care providers [27]. Moreover, it improves patient well-being and helps caregivers make informed clinical decisions [28]. In fact, the use of such apps not only benefits patients while receiving treatment but also provides tools such as follow-up care and self-management for breast cancer survivors [29]. Patients’ need for self-management techniques is crucial, as it helps them to make their lives better by complying with the treatment needed and, as a result, accepting the disease [30].

However, despite the high number of downloads and star ratings of health apps, including those specifically targeting breast cancer, several challenges remain with respect to their quality and security. Evidence from the literature reports on the existence of poor-quality health apps that fall short with respect to (1) following evidence-based health guidelines and best practices, (2) involving experts and consumers in their development, and (3) demonstrating effectiveness based on empirical evidence, all of which ultimately can be potentially harmful to their users [14,31-33].

Additionally, health apps have been facing critical challenges related to their privacy, confidentiality, and security [14,34,35], especially given their nature of handling sensitive, personal, and health-related data [36]. These challenges have been magnified with the rise of cyberattacks through apps and mobile devices [37] and further highlighted by recent regulations such as the General Data Protection Regulation for member states of the European Union [38]. Such a regulation assesses the privacy score of mobile apps and identifies or measures apps’ privacy based on 14 components [39].

As the uptake of these apps increases, it becomes imperative for users to evaluate their quality and safety [40]. Despite the high prevalence of breast cancer among the population of the MENA region, evidence regarding the quality, privacy, and security of breast cancer apps available to Arabic speakers remains poor. This study aimed to conduct a systematic assessment of mobile breast cancer apps available for Arabic speakers to evaluate their functionality, quality, security, and data safety. To the best of our knowledge, no previous study has addressed this gap.

Methods
Overview
Using a similar approach to several recent studies [41-45] and to ensure scientific rigor, this study conducted a systematic search and content analysis of mobile breast cancer apps available for Arabic speakers. We searched both Google Play and Apple App stores between December 18 and 24, 2022.

Search Strategy
Initially, we used the Arabic search terms highlighted in Table 1 to search Google Play and Apple App stores. These search terms were selected based on earlier work that was published concerning the top searched breast cancer topics by Arabic speakers using Google’s search engine [8]. To ensure rigor, the researchers searched the app stores both through (1) the web interface and (2) natively on devices running the relevant operating system, thus mimicking how end users will discover such apps.
Table 1. The terms used to search for breast cancer apps available to Arabic speakers and their English translations.

<table>
<thead>
<tr>
<th>Arabic term</th>
<th>Translated term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer</td>
<td>Breast cancer</td>
</tr>
<tr>
<td>Breast cancer screening</td>
<td>Breast cancer screening</td>
</tr>
<tr>
<td>Breast cancer treatment or therapy</td>
<td>Breast cancer treatment or therapy</td>
</tr>
<tr>
<td>Breast cancer detection or screening</td>
<td>Breast cancer detection or screening</td>
</tr>
<tr>
<td>Breast cancer disease</td>
<td>Breast cancer disease</td>
</tr>
<tr>
<td>Symptoms of breast cancer</td>
<td>Symptoms of breast cancer</td>
</tr>
<tr>
<td>Breast screening</td>
<td>Breast screening</td>
</tr>
</tbody>
</table>

Eligibility

App eligibility was determined by 2 independent researchers blinded to each other’s decisions, and the apps were initially screened based on the app’s name, the provided screenshots, and the app’s description. Discrepancies between researchers were resolved through consensus. Apps were included if they were free of charge, provided content and support for Arabic speakers, and were designed for use by consumers or patients; all apps were considered regardless of release or last update dates (Textbox 1). Apps were excluded if they were paid or were subscription-based, did not support Arabic speakers, or were designed for use by clinicians or health care workers.

Textbox 1. Inclusion and exclusion criteria for the apps.

**Inclusion criteria**

- Free of charge
- Available on Google Play or Apple App store
- Designed for use by consumers or patients
- Support Arabic speakers
- Considered regardless of release or last update dates

**Exclusion criteria**

- Paid or subscription based
- Designed for use by clinicians or health care workers
- Does not support Arabic speakers

Data Extraction and Evaluation

Initially, all information provided by the app developers in the app stores was extracted to evaluate the descriptive features and the general characteristics of the included apps, which included the platform, developer name, update date, ratings, number of reviews, number of downloads, app category, and app permissions, as reported by the app developers. Afterward, 2 independent researchers downloaded the apps on their smartphones to assess the quality and privacy risks of the included apps.

We evaluated the quality of the included apps using a standardized form, the Mobile App Rating Scale (MARS), focusing on the following 4 dimensions: engagement, functionality, aesthetics, and information quality [46,47]. All scores were compared among 2 researchers, and the average score for each dimension was reported. To evaluate the apps’ privacy risks, we assigned scores to the permissions requested by the apps as reported by the app developers. The scores were informed by previous research, where the score risk is 0 for nonthreatening, 0.5 for potentially threatening, and 1 for threatening permissions [48]. Such permissions include access to restricted data, such as system state and user contact information, and restricted actions, such as connecting to a paired device and recording audio [49]. The 2 researchers independently carried out this evaluation and were unaware of each other’s scores; any discrepancies were resolved through consensus.

Results

Overview

The researchers followed the systemic steps, highlighted in Figure 1, resulting in the inclusion of 9 apps, all of which are Android apps found on the Google Play store.

Overall, the included apps were indicated to be appropriate for all ages and were either in the medical, education, lifestyle, personalization, or health and fitness categories as per Google Play store categorization (Table 2). Our investigation suggests that the apps were all informational in nature, mainly providing information about breast cancer. None of the included apps had a language option to make it available in more than 1 language.
At the time of data collection, the results show that the included apps had low overall downloads (>10 to >1000) and more than half (6/9, 67%) were supported by advertisements. Only 5 apps had reviews, with an average of 10.60 reviews and an average star rating of 4.78. Only 1 app was last updated in 2019, while the remaining apps were updated in the past 2 years.

**Figure 1.** The flow diagram of the systematic search process to identify the relevant apps.

![Flow Diagram](image)

**Table 2.** Characteristics of the included breast cancer apps in the Google Play store.

<table>
<thead>
<tr>
<th>App number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Stars</th>
<th>Reviews, n</th>
<th>Advertisement supported</th>
<th>Downloads, n</th>
<th>Rating&lt;sup&gt;b,c&lt;/sup&gt;</th>
<th>Updated on&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Category&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>No</td>
<td>&gt;100</td>
<td>3+</td>
<td>May 11, 2021</td>
<td>Medical</td>
</tr>
<tr>
<td>2</td>
<td>4.9</td>
<td>23</td>
<td>Yes</td>
<td>&gt;1000</td>
<td>3+</td>
<td>July 26, 2019</td>
<td>Medical</td>
</tr>
<tr>
<td>3</td>
<td>_</td>
<td>_</td>
<td>Yes</td>
<td>&gt;100</td>
<td>3+</td>
<td>April 18, 2020</td>
<td>Health and fitness</td>
</tr>
<tr>
<td>4</td>
<td>4.2</td>
<td>6</td>
<td>Yes</td>
<td>&gt;1000</td>
<td>3+</td>
<td>November 3, 2020</td>
<td>Education</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>8</td>
<td>No</td>
<td>&gt;500</td>
<td>3+</td>
<td>May 11, 2021</td>
<td>Medical</td>
</tr>
<tr>
<td>6</td>
<td>4.8</td>
<td>6</td>
<td>No</td>
<td>50</td>
<td>3+</td>
<td>July 11, 2022</td>
<td>Medical</td>
</tr>
<tr>
<td>7</td>
<td>_</td>
<td>_</td>
<td>Yes</td>
<td>500</td>
<td>3+</td>
<td>August 9, 2021</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>8</td>
<td>_</td>
<td>_</td>
<td>Yes</td>
<td>10</td>
<td>3+</td>
<td>October 8, 2022</td>
<td>Personalization</td>
</tr>
<tr>
<td>9</td>
<td>_</td>
<td>_</td>
<td>Yes</td>
<td>10</td>
<td>3+</td>
<td>October 8, 2022</td>
<td>Personalization</td>
</tr>
</tbody>
</table>

<sup>a</sup>Arbitrary number to mask app name.

<sup>b</sup>Based on Google Play store.

<sup>c</sup>Content of apps considered suitable for age group indicated per Google Play store rating.

<sup>d</sup>At data collection.

<sup>e</sup>Not available.
App Permissions and Controls

As described in the Google Play store, the majority of the included apps (8/9, 89%) were requesting dangerous access permissions, including access to storage, media files, and camera permissions (Table 3). Additional permissions were exposed after the researchers downloaded the apps on the testing devices. These permissions included receiving data from the internet; viewing network connections; having full network access; running at start-up; controlling flashlight and vibration; preventing the device from sleeping; reading badge notifications; running foreground services; advertising permissions; reading location from media; playing and installing referrer application programming interface; and lastly, pairing with Bluetooth devices (Table S1 in Multimedia Appendix 1). The researchers considered all permissions and added them up to calculate the final risk score (Table 4).

On average, the security score for the included apps is 3.22 (total points possible: 9.5). The security scores ranged between 0 and 7.5; only 2 apps had a total security score of 0. The apps requested potentially dangerous permissions, namely full network access, advertising ID permission, read location from media collection, precise location (GPS and network-based), take pictures and videos, read the contents of the USB storage, modify or delete the contents of the USB storage, and view Wi-Fi connections.

Table 3. Included apps’ permissions as described in the Google Play store.

<table>
<thead>
<tr>
<th>App number</th>
<th>Location</th>
<th>Camera</th>
<th>Photos and media</th>
<th>Storage</th>
<th>Wi-Fi connection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Approximate location (network based)</td>
<td>Precise location (GPS and network based)</td>
<td>Take pictures and videos</td>
<td>Read the contents of your USB storage</td>
<td>Modify or delete the contents of your USB storage</td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 4. The security score per app based on its permissions requested or declared.

<table>
<thead>
<tr>
<th>Permission</th>
<th>Weight&lt;sup&gt;a&lt;/sup&gt;</th>
<th>App 1</th>
<th>App 2</th>
<th>App 3</th>
<th>App 4</th>
<th>App 5</th>
<th>App 6</th>
<th>App 7</th>
<th>App 8</th>
<th>App 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receive data from the internet</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>View network connections</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Full network access</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Run at start-up</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Control flashlight</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Control vibration</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Prevent device from sleeping</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Read badge notification</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Run foreground service</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Advertising ID permission</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Read location from media collection</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Play install referrer API&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Pair with Bluetooth devices</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Approximate location (network-based)</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Precise location (GPS and network-based)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Take pictures and videos</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Modify or delete the contents of your USB storage</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Read the contents of your USB storage</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Modify or delete the contents of your USB storage</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>View Wi-Fi connections</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total security score per app</td>
<td>9.5</td>
<td>0</td>
<td>2.5</td>
<td>7.5</td>
<td>2.0</td>
<td>6.0</td>
<td>5.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

<sup>a</sup>The threat weight was calculated following the guidance provided by Olmstead and Atkinson [48].

<sup>b</sup>API: application programming interface.

### Data Safety

Only 3 apps provided information about data safety and how the data are handled (Table S2 in Multimedia Appendix 1). Specifically, only 2 apps reported not sharing user data with other companies or organizations; the same apps reported that their apps do not collect user data. On the other hand, only 1 reported sharing information such as location, health and fitness messages, as well as photos and videos. The 3 apps that declared information about data safety reported that the data are encrypted in transit. Only 1 app reported that the users can request to delete the data.

### Apps’ Quality Rating

The researchers used MARS to evaluate the included apps’ quality. MARS uses 4 dimensions to assess the apps: engagement, functionality, aesthetics, and information. The average overall score rating of the included apps was 3.27 (Table 5). Notably, the evaluation showed that all the apps had a high rate in the functionality dimension, where they all scored 4.75; as for the engagement dimension, all apps had a meager score with an average of 2.32. Considering the other 2 dimensions, aesthetics and information, we can see that the scores vary between apps, where some have a high score and others have a low score, with an average of 3.00 and 3.04, respectively.
Discussion

Principal Findings

To the best of our knowledge, this work summarizes the most extensive collection of the currently available free-of-charge breast cancer apps for Arabic speakers on the Google Play and Apple App stores. Interestingly, this systematic investigation reveals that, at the time of data collection, no breast cancer apps were identified as available to Arabic speakers in the Apple App store. Additionally, none of the apps at the time of data collection provided language options, which can hinder the possibility of translating to multiple languages. Overall, the results of this research showed a lack of breast cancer apps available to Arabic speakers compared to apps available in Turkish [50] and Korean [51] for example.

The analyzed apps in this study are available for free, with the majority of these apps (6/9, 67%) being supported by advertisements. The apps varied in what they are able to access and control on the devices, as shown in Table 3. A total of 44% (4/9) apps use Wi-Fi, 33% (3/9) can access the device’s storage, 22% (2/9) have access to the device’s camera, and 11% (1/9) are able to access the media as well as the location. Only 33% (3/9) of the apps stated how the collected data would be used (Table S2 in Multimedia Appendix 1), while the rest of the apps did not specify any information.

The results of this study demonstrate that the quality of the considered apps is highly “functional” but less “engaging” (Table 5). The average overall score rating of the included apps (3.27) is slightly above average, suggesting that the majority of the apps may not be considered exceptional by consumers. While there seem to be positive aspects to the included apps, there is still room for improvement. Considering Arabic content related to breast cancer, recent evidence suggests the low quality of informational videos available on YouTube despite the high number of views [9]. Our findings provide evidence suggesting that nearly average-quality information content about breast cancer is available to Arabic speakers.

A recent Spanish study tested 6 apps (2 on iOS, 5 on Android, and 1 on both) using the MARS framework. The results of the study showed relatively diverse scores, with an objective quality mean score of 3.06 and a subjective quality mean score of 1.96 [52]. Another study evaluated the quality of mHealth apps for educational purposes in Iran using the MARS framework. The study reported a mean score of 4.01 for quality and 3.08 for subjective quality [53]. Turkish apps were also evaluated using the MARS framework, and the study found an average score of 3.31 [50], which is similar to the Spanish study’s results.

Although many research papers have addressed the importance of using mHealth apps to improve patients’ health, provide educational materials, enhance communication between patients and caregivers, and achieve a successful recovery, these studies have stated that such apps are exposed to several challenges and threats as well. For instance, middle- and high-income households had more access to the internet on their mobile devices compared to those with low incomes [28]. Therefore, patients and caregivers who have no or limited access to the internet may not be able to use the app’s features effectively, or they might not prefer to use a mobile phone for health-related purposes in general.

Another concern is that the process of storing and transferring personal health data through a mobile app could be insecure and might cause serious security and privacy issues [28,54]. Thus, mobile apps should be designed to accommodate a wide range of possible users while considering health knowledge, different levels of cultural needs, and linguistic requirements. Additionally, recent research suggests that assessing the apps’ quality, safety, and usability by involving patients and health care professionals will result in an ideal app that meets patients’ requirements and enhances the app’s overall safety as well [20].

The literature on breast cancer smartphone apps suggests that these apps can be a valuable resource of information for individuals with the disease [55]. These apps provide information on treatment options, support groups, and local resources. The apps also allow individuals to track their symptoms, set reminders for appointments, and record their progress throughout treatment [56]. Such features can significantly contribute to enhancing patients’ well-being [57]. Additionally, studies have found that these apps can improve communication with health care providers and improve self-efficacy and quality of life [30,58]. However, it is important to note that the quality and accuracy of the information provided

### Table 5. Mobile App Rating Scale evaluation for Arabic breast cancer apps.

<table>
<thead>
<tr>
<th>App number</th>
<th>Engagement</th>
<th>Functionality</th>
<th>Aesthetics</th>
<th>Information</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.30</td>
<td>4.75</td>
<td>3.67</td>
<td>3.87</td>
<td>3.60</td>
</tr>
<tr>
<td>2</td>
<td>2.40</td>
<td>4.75</td>
<td>3.00</td>
<td>3.00</td>
<td>3.30</td>
</tr>
<tr>
<td>3</td>
<td>2.40</td>
<td>4.75</td>
<td>3.00</td>
<td>2.25</td>
<td>3.10</td>
</tr>
<tr>
<td>4</td>
<td>2.40</td>
<td>4.75</td>
<td>3.33</td>
<td>2.50</td>
<td>3.20</td>
</tr>
<tr>
<td>5</td>
<td>2.40</td>
<td>4.75</td>
<td>3.33</td>
<td>3.50</td>
<td>3.50</td>
</tr>
<tr>
<td>6</td>
<td>2.40</td>
<td>4.75</td>
<td>3.67</td>
<td>4.00</td>
<td>3.70</td>
</tr>
<tr>
<td>7</td>
<td>2.20</td>
<td>4.75</td>
<td>3.00</td>
<td>2.75</td>
<td>3.20</td>
</tr>
<tr>
<td>8</td>
<td>2.20</td>
<td>4.75</td>
<td>2.00</td>
<td>2.75</td>
<td>2.90</td>
</tr>
<tr>
<td>9</td>
<td>2.20</td>
<td>4.75</td>
<td>2.00</td>
<td>2.75</td>
<td>2.90</td>
</tr>
<tr>
<td>Total mean score</td>
<td>2.32</td>
<td>4.75</td>
<td>3.00</td>
<td>3.04</td>
<td>3.27</td>
</tr>
</tbody>
</table>

https://cancer.jmir.org/2024/1/e48428
by these apps can vary, and it is recommended to consult with a health care professional before using any app for managing a medical condition.

A recent study on breast health and breast cancer apps notes that although apps appear to be competitive and useful for patients, some major features have to be considered while developing these apps [59]. The features include notifications, reminders, symptoms tracking, and recording. The study also suggests designing the apps to be user-friendly, even for low-literacy patients. Developing features with audio support will not only help patients with low literacy but can also support multiple languages.

**Recommendations and Implications for Practice**

**Privacy and Security**

We would recommend that the developers of the apps be more transparent and state how the data will be used and that they should not have access to unnecessary data. We recommend that future breast cancer apps be available to Arabic speakers to justify the need for the permissions requested while also transparently disclosing the data safety handling measures to the app users. Security and privacy of apps are considered major requirements as they are accountable for sensitive patient data such as prescriptions, treatments, etc. Thus, to come up with robust apps that could ensure privacy and security appropriately, more evaluation techniques, as well as security mechanisms, should be analyzed and implemented on Arabic apps, in particular, to assess, measure, and control the apps’ security and privacy [60].

**Quality and Engagement**

Involving patients and health professionals in the app design phase is crucial. Several studies have addressed the idea that health applications should be developed and designed based on the combined efforts of health professionals, related academics, and patients [61]. To raise the quality of breast cancer apps, the inclusion of utility features such as appointment booking for mammograms and web-based consultations becomes necessary. In addition, it is recommended to improve health apps’ engagement by focusing on specific components such as personalized content, data visualization, reminders and notifications, educational material, self-management functions, and goal-setting features [62]. Providing users and patients with proper communication features and a well-designed interface leads to an ideal user experience as a result [30].

**Study Strengths and Limitations**

Similar to other studies, a rigorous multistep methodology mimicking systematic reviews is used in this study to assess the breast cancer apps that are available to Arabic speakers. Apps were thoroughly searched through both the web interface as well as the app stores natively on the devices, mimicking how end users will discover such apps. While the results provide an indication of the quality of the evaluated apps, additional investigations are required to consider patients’ perspectives about their views about the quality as well as the utility of such apps. Future studies can also involve rigorous assessments with respect to the security measures applied by breast cancer apps available to Arabic speakers.

This study only considered the publicly available apps and may have missed apps that are “prescribed” to patients or consumers or those that are developed locally by health care organizations. Another limitation, which is inherent to the search strategy used in this work as well as similar other work [63], is the fact that the search algorithms used by the app stores are nontransparent and can change without the public’s knowledge, potentially undermining the reproducibility of the outcomes. Lastly, the current state of the results as revealed by this work is likely to change quite rapidly since apps are regularly released, updated, and retired.

**Conclusions**

The battle against breast cancer is not over yet, and breast cancer apps can serve as valuable resources in this ongoing fight. The results of this systematic and thorough examination of breast cancer apps available for Arabic speakers reveal their limited existence at the time of study. The investigations evaluated these apps through the lenses of quality, privacy, and security, revealing that the included apps are rated as highly “functional” but at the same time are less “engaging.” The investigations also reveal that some apps were accessing unnecessary data and collecting information that was not relevant to the purpose of the app.

Developers of breast cancer apps that cater to Arabic speakers must focus on consumers’ preferences, demographics, usability, and the interface of their apps, as well as enhance measures and mechanisms of privacy and security for their apps. The low number of breast cancer apps available to Arabic speakers, as revealed in this study, should be a call to action for many health care organizations and developers to collaboratively develop information-rich, usable, functional, engaging, and secure apps.

**Data Availability**

All data generated or analyzed during this study are included in this paper and its supplementary information files.

**Authors’ Contributions**

DA conceptualized the study design and supervised the study. DA and AA participated in data assembly, analysis, and interpretation, and in the writing of the paper. LA contributed to data interpretation and in the writing of the paper. All authors read and approved the final paper.
Conflicts of Interest

None declared.

Multimedia Appendix 1

Included apps' declared permission and safety declarations.

References


Abbreviations

MARS: Mobile App Rating Scale
MENA: Middle East and North Africa
mHealth: mobile health
Assessing the Quality, Privacy, and Security of Breast Cancer Apps for Arabic Speakers: Systematic Search and Review of Smartphone Apps

Alhuwail D, Alhouti A, Alsarhan L

JMIR Cancer 2024;10:e48428
URL: https://cancer.jmir.org/2024/1/e48428
doi: 10.2196/48428
PMID: 38227353

© Dari Alhuwail, Aisha Alhouti, Latifah Alsarhan. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Comparison of the Real-World Reporting of Symptoms and Well-Being for the HER2-Directed Trastuzumab Biosimilar Ogivri With Registry Data for Herceptin in the Treatment of Breast Cancer: Prospective Observational Study (OGIPRO) of Electronic Patient-Reported Outcomes

Andreas Trojan1,2, MD; Sven Roth3, BSc; Ziad Atassi2, MD; Michael Kiessling2, MD, PhD; Reinhard Zenhaeusern4, MD; Yannick Kadvany5, MA; Johannes Schumacher6, PhD; Gerd A Kullak-Ublick1, Prof Dr Med; Matti Aapro7, MD; Alexandru Eniu8, MD

1Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, University of Zurich, Zurich, Switzerland
2BrustZentrum Zürichsee, Horgen, Switzerland
3Faculty of Medicine, University of Zurich, Zurich, Switzerland
4Onkologie, Spital Oberwallis, Brig, Switzerland
5Mobile Health AG, Zurich, Switzerland
6Palleos Healthcare, Wiesbaden, Germany
7Cancer Center, Clinique de Genolier, Genolier, Switzerland
8Hôpital Riviera-Chablais, Rennaz, Switzerland

Corresponding Author:
Andreas Trojan, MD
Department of Clinical Pharmacology and Toxicology
University Hospital Zurich
University of Zurich
Rämistrasse 100
Zurich, 8091
Switzerland
Phone: 41 76 34 30 200
Email: trojan@1st.ch

Abstract

Background: Trastuzumab has had a major impact on the treatment of human epidermal growth factor receptor 2 (HER2)-positive breast cancer (BC). Anti-HER2 biosimilars such as Ogivri have demonstrated safety and clinical equivalence to trastuzumab (using Herceptin as the reference product) in clinical trials. To our knowledge, there has been no real-world report of the side effects and quality of life (QoL) in patients treated with biosimilars using electronic patient-reported outcomes (ePROs).

Objective: The primary objective of this prospective observational study (OGIPRO study) was to compare the ePRO data related to treatment side effects collected with the medidux app in patients with HER2-positive BC treated with the trastuzumab biosimilar Ogivri (prospective cohort) to those obtained from historical cohorts treated with Herceptin alone or combined with pertuzumab and/or chemotherapy (ClinicalTrials.gov NCT02004496 and NCT03578731).

Methods: Patients were treated with Ogivri alone or combined with pertuzumab and/or chemotherapy and hormone therapy in (neo)adjuvant and palliative settings. Patients used the medidux app to dynamically record symptoms (according to the Common Terminology Criteria for Adverse Events [CTCAE]), well-being (according to the Eastern Cooperative Oncology Group Performance Status scale), QoL (using the EQ-5D-5L questionnaire), cognitive capabilities, and vital parameters over 6 weeks. The primary endpoint was the mean CTCAE score. Key secondary endpoints included the mean well-being score. Data of this prospective cohort were compared with those of the historical cohorts (n=38 patients; median age 51, range 31-78 years).

Results: Overall, 53 female patients with a median age of 54 years (range 31-87 years) were enrolled in the OGIPRO study. The mean CTCAE score was analyzed in 50 patients with available data on symptoms, while the mean well-being score was evaluated in 52 patients with available data. The most common symptoms reported in both cohorts included fatigue, taste disorder,
Biosimilars and reference biologics play a key role in the treatment of cancer and account for approximately 70% of the growth in costs of drugs from 2010 to 2015 [1]. Therefore, pricing is an important challenge for the medical society and biosimilars offer an attractive option for a value-based care environment with cost-saving potential [2].

Trastuzumab (Herceptin), a human epidermal growth factor receptor 2 (HER2) antibody, has had a major impact on the treatment of patients with HER2-positive breast cancer (BC) worldwide, which now has indications for the treatment of small tumors in both (neo)adjuvant and palliative settings [3,4]. This provides a good opportunity to compare the efficacy and safety of trastuzumab biosimilars to those of trastuzumab in clinical trials. For several anti-HER2 biosimilars, safety and clinical equivalence to the reference product have been demonstrated [2,5]. In a randomized, parallel-group phase 3 equivalence study of patients with HER2-positive metastatic BC, Rugo et al [6] demonstrated equivalent efficacy and similar safety profiles between the trastuzumab biosimilar Ogivri (MYL-1401O) and trastuzumab (Herceptin) [6].

The enhanced assessment of electronic patient-reported outcomes (ePROs) in clinical routine and cancer trials is of growing interest [7-9]. Several studies indicate that the proactive use of PROs can identify otherwise undetected symptoms and improve symptom management for patients with various types of cancer [9] as well as offer improvements in well-being and awareness of adverse events (AEs) between outpatient visits. Using a mobile app, especially in collaboration with the treating physician, might improve clinical care in patients with early or advanced disease [10-13]. In addition, the benefits of digital patient monitoring have been demonstrated during immune and targeted cancer therapies in terms of more efficient symptom assessment and patient-physician communication as well as a reduced need for telephone consultations [14].

Medidux is an interactive patient empowerment app that enables physicians, especially oncologists, to monitor the progression of well-being and symptoms of patients undergoing cancer treatment. Based on the documented symptom progression, the software notifies patients to contact the treatment team if symptoms defined according to the Common Terminology Criteria for Adverse Events (CTCAE) standards are outside the acceptable range. More than 110 available symptoms and severity classifications (according to the CTCAE), as well as high numbers of standardized symptom reports from patients, contribute to the collection of high-quality ePRO data for the timely management of treatment-related AEs and toxicities and their communication to treatment teams [11,13]. Thus, the medidux smartphone app is helpful to stabilize daily functional activities and leads to more frequent reporting of AEs and more precise entries regarding symptoms [11]. The continuous measurement of ePROs enables structured and standardized data recording of patients’ daily health state.

An increased level of concordance (κ=0.68) for common symptoms, including pain, fever, diarrhea, constipation, nausea, vomiting, and stomatitis, between the patient and treating physician was recently demonstrated for the medidux platform [13]. However, to the best of our knowledge, no real-world observation of side effects, tolerability, and quality of life (QoL) has been performed using ePRO data collected from patients treated with anti-HER2 biosimilars. Thus, the aim of this observational study was to investigate real-world data on daily functional activity, symptoms, and therapy side effects recorded with the medidux smartphone app in patients undergoing Ogivri antibody therapy. In addition, historical ePRO data of 38 patients with HER2-positive BC treated with Herceptin from two previous studies [7,13] were used for comparative analysis.

### Methods

#### Study Design

OGIPRO was a noninterventional, multicenter, prospective, observational study conducted at 5 study sites in Switzerland over a duration of 20 months.

Patients 18 years and older with a histologically or cytologically proven diagnosis of HER2-positive primary, locally advanced, or metastatic BC were eligible to participate after providing written informed consent. In addition, patients had to own a personal iOS or Android smartphone.
Eligible patients received anti-HER2 treatment containing the trastuzumab biosimilar Ogivri (initial dose of 8 mg/kg body weight [BW] intravenously, followed by 6 mg/kg BW) with or without pertuzumab and/or chemotherapy and hormone therapy in (neo)adjuvant and palliative settings. At the beginning of the study, patients were provided with the medidux app and were prompted to record their symptoms, well-being, EQ-5D-5L questionnaires, cognitive capabilities, and vital parameters every day. Patients underwent 3 regular study visits scheduled on days 1, 21, and 42 during their 3 weekly chemotherapeutic interventions. All anticancer treatments used in this study were approved drugs, and the therapy was compliant with national treatment guidelines.

The observational period for each patient was 6 weeks. At the end of the observational period, patients decided whether to continue their therapy with the biosimilar Ogivri or with the reference substance Herceptin.

After the study observational period, prospectively collected data of patients treated with Ogivri (prospective cohort) were compared to historical ePRO data of patients treated with Herceptin (historical cohort) in two previous studies: a prospective randomized controlled trial (PRO1 study; ClinicalTrials.gov NCT02004496) of 139 patients with early stage BC who underwent chemotherapy [7] and an observational noninterventional trial (Consilium1 study; ClinicalTrials.gov NCT03349713) of patients with breast, colon, prostate, or lung cancer undergoing cancer treatment [13]. In both studies, patients were encouraged to document data on well-being and standardized symptoms using earlier versions of the medidux app during the course of their therapies. More than 5000 continuously measured data entries from 38 patients overall (14 from Consilium1 and 24 from PRO1) were available for the comparative analysis [7,13]. The historical ePRO data were recorded in the same manner using the earlier versions of the mobile app [11] and were therefore comparable to the prospective ePRO data.

**Ethical Considerations**

This study was approved by the Swiss Institutional Review Board (KEK-ZH: 2021-D0051) and was conducted in accordance with the principles of the Declaration of Helsinki (current version). The study was also registered on ClinicalTrials.gov (NCT05234021). All patients in the prospective and historical cohorts provided written informed consent prior to enrollment and were informed that participation in the study is voluntary and can be revoked at any time. All study documents were deidentified by assigning a unique ID to each patient. Functional data security was ensured by identification only made possible via the patient’s ID. The data on the patients’ devices were encrypted in the app and the data exchange was encrypted with the patient’s ID. There was no compensation provided to participants.

**Objectives**

**Primary Objective**

The primary objective of the study was to evaluate ePRO data reported in the medidux app by patients with HER2-positive BC treated with the trastuzumab biosimilar Ogivri with respect to their treatment side effects and to compare these data with ePRO data obtained from a historical cohort of 38 patients treated with Herceptin in two previous studies (NCT02004496, NCT03349713) [7,11,13]. No difference was expected for the CTCAE score between the two cohorts. The aims of the study were therefore to confirm that the average CTCAE scores were similar in both cohorts and that the recording of side effects with the app was reliable.

**Secondary Objectives**

Secondary objectives included well-being in both cohorts as well as electronically reported symptoms with respect to the therapy regimen and demographic characteristics only in the prospective cohort.

**Mobile App**

The medidux app (version 3.2) used in the study is a patient-centered, therapy-accompanying app that supports the structured, standardized, and dynamic documentation of symptoms and therapy side effects. Use of this tool does not represent an invasive intervention on the patient and consequently did not pose any specific risks of physical injury.

**Data Collection**

The app has two basic components: (1) a browser-based app for the treatment team (web app) and (2) a mobile app for cancer patients. There was no need for 24-hour monitoring by medical staff in connection with use of the app.

The medidux app for patients enabled recording symptoms, vital signs, and well-being in a structured and standardized manner. Patients were encouraged to regularly enter data on symptoms according to the CTCAE (version 4.0), general well-being according to the Eastern Cooperative Oncology Group Performance Status (ECOG PS), EQ-5D-5L questionnaire (weekly), vital signs (weight, blood glucose, blood pressure, and pulse), and optionally a neuropsychological cognitive test (Trail Making Test [TMT]), concomitant medications, and private notes. Patients were asked daily about their general well-being and symptoms using a visual analog scale. Recording usually started on the day of therapy initiation (or the change in therapy) and continued through an observational period of 6 weeks. The frequency of app use and data entry was logged throughout the course of the study treatment, which served as an indicator of patients’ active participation in the study and as a relevant process parameter for evaluating the usability of the app itself.

The mobile app also recommended contacting the investigator or treatment site in case of high intensity of symptoms (ie, treatment side effects). Furthermore, the app provided patients with self-efficacious recommendations and tips on how to treat and reduce treatment side effects.

**Recording of AEs**

AEs in the app were classified according to the CTCAE (version 4.0). For the app, grade 5 “Death related to AE” had been removed. Instead, category 0 was added, representing no or very mild symptoms. The 5 severity levels were translated into a visual analog scale from 0.1 to 10, with 0.1 representing the weakest possible symptom and 10 representing the strongest.
possible symptom. Scores 0.1-2.0 corresponded to grade 0, scores 2.1-4.0 corresponded to grade 1, scores 4.1-6 corresponded to grade 2, scores 6.1-8 corresponded to grade 3, and scores 8.1-10 corresponded to grade 4 AEs. When patients selected a score between 0.1 and 10, they received a summary and information for the selected range, which was displayed in the app. Classification into adapted grades based on the CTCAE resulted in the following categories: minimal symptoms (0), mild symptoms (1), moderate symptoms (2), severe symptoms (3), and very severe symptoms (4).

Well-Being Assessment
Self-assessment of well-being was carried out in the medidux app with the help of a slider on a visual analog scale that allows for the continuous selection from 0 to 100. At the same time, short definitions appear for the standardized and structured reporting of the gradings, which should help the patient to correctly categorize their well-being. Selected values between 81 and 100 correspond to an ECOG PS of 0, values of 61-80 correspond to ECOG PS 1, values of 41-60 correspond to ECOG PS 2, values of 21-40 correspond to ECOG PS 3, and values of 0-20 correspond to ECOG PS 4. As mentioned above, grade 5 “Dead” was removed for the app.

Statistical Analyses

Sample Size Calculation
The research objective was to investigate the difference between prospective and historical cohorts regarding patient-reported side effects, operationalized by the CTCAE score over a period of 6 weeks. To assess the hypothesis of equal CTCAE scores in both cohorts, the method of interval estimation was selected using the 95% CI for the mean difference between cohorts. A statistical analysis plan (SAP) prospectively determined the required sample size for a prospective cohort based on the data from the historical cohort (as available from the previous studies NCT02004496 and NCT03578731 [7,11,13]; see the Study Design section above for further details). First, the SD for the CTCAE scores of the 38 patients in the historical cohort was calculated retrospectively as 9.7 and the assumption of an equal SD in the prospective cohort was made. Second, the sample size for the prospective cohort was chosen to achieve a certain minimum precision in estimating the mean difference between cohorts (width of the 95% CI). For a range of feasible sample sizes, the SAP reported 95% CI precisions based on the t distribution (calculated using the R package presize [15]), assuming an equal SD of 9.7 in both cohorts and using a pooled variance estimate. From this range, a sample size of 60 patients was prospectively selected in the SAP to achieve 51 evaluable patients, given an expected dropout rate of 15%. The corresponding 95% CI for the mean difference between the historical and prospective cohorts was estimated to have a precision of 8.3, which was deemed acceptable for the planned assessment in the given study context.

Statistical Methods
All analyses of the primary and secondary endpoints (CTCAE score, well-being score) were performed using univariate analyses, followed by multivariate linear regression to report (adjusted) mean differences between historical and prospective cohorts, with the P values based on t tests and corresponding 95% CIs. All multivariate models extended the respective univariate models in a supplementary fashion to adjust for potential imbalances in patient age, tumor stage, and therapy setting. These covariates were prospectively defined in the study’s SAP; no model selection procedures were employed. All analyses were performed using R version 4.2.0 (The R Foundation for Statistical Computing) [16]. Two-sided P values ≤.05 were considered statistically significant.

Primary Endpoint
The primary endpoint, a CTCAE score based on the severity grades of the 10 most relevant side effects (sensory disturbance, diarrhea, fatigue, nausea, vomiting, headache, fever, edema of the limbs, joint pain, and loss of appetite) after 6 weeks, was compared between the prospective and historical cohorts. The CTCAE score was calculated by averaging the score per patient and symptom and then averaging the score per patient over all symptoms. The mean difference in the CTCAE scores between cohorts was estimated using univariate linear regression with 95% CIs. To adjust for potential differences between the two cohorts in covariates relevant for the primary outcome, a supplementary multivariate analysis was performed including the additional covariates patient age, tumor stage, and therapy setting.

Secondary Endpoint
The well-being score according to the ECOG PS was collected continuously using a visual analog scale (range 0-100) implemented in the medidux app and averaged across measurements during the observational period. The analysis protocol was analogous to that described above for the primary endpoint.

Additional Analysis
Cognitive tests in the prospective cohort were collected continuously throughout the observation period and descriptively assessed by administering a modified version of the TMT. The time (in seconds) to complete each task (execution time) was used in the analysis.

Results

Baseline Characteristics
Overall, 53 female patients were enrolled in the OGIPRO study. The median age was 57 (range 34-87) years in the prospective cohort and 51 (range 31-78) years in the historical cohort. Most patients (38.9%) had tumor stage 2 (Table 1). With regard to the treatment setting, relatively fewer patients (22.2%) received palliative treatment than neoadjuvant or adjuvant treatment. More than half of the patients (59.3%) received dual anti-HER2 blockade with trastuzumab and pertuzumab (Table 1).
Table 1. Baseline characteristics.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Historical cohort (n=38)</th>
<th>Prospective cohort (n=53)</th>
<th>Total (N=91)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.001a</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>51.3 (10)b</td>
<td>59.09 (12.193)</td>
<td>55.89 (11.924)b</td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>51 (31-78)</td>
<td>57 (34-87)</td>
<td>54 (31-87)</td>
<td></td>
</tr>
<tr>
<td><strong>Tumor stage (T), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.07c</td>
</tr>
<tr>
<td>T1</td>
<td>8 (21.05)</td>
<td>10 (19.23)b</td>
<td>18 (20.0)b</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>19 (50.0)</td>
<td>16 (30.77)b</td>
<td>35 (38.89)b</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>8 (21.05)</td>
<td>11 (21.15)b</td>
<td>19 (21.11)b</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>3 (7.89)</td>
<td>15 (28.85)b</td>
<td>18 (20.0)b</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment setting, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.02c</td>
</tr>
<tr>
<td>Neoadjuvant</td>
<td>17 (44.74)</td>
<td>18 (34.62)b</td>
<td>35 (38.89)b</td>
<td></td>
</tr>
<tr>
<td>Adjuvant</td>
<td>18 (47.37)</td>
<td>17 (32.69)b</td>
<td>35 (38.89)b</td>
<td></td>
</tr>
<tr>
<td>Palliative</td>
<td>3 (7.89)</td>
<td>17 (32.69)b</td>
<td>20 (22.22)b</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.01c</td>
</tr>
<tr>
<td>Trastuzumab</td>
<td>1 (2.63)</td>
<td>10 (18.87)</td>
<td>11 (12.09)</td>
<td></td>
</tr>
<tr>
<td>Trastuzumab+pertuzumab</td>
<td>21 (55.26)</td>
<td>33 (62.26)</td>
<td>54 (59.34)</td>
<td></td>
</tr>
<tr>
<td>Trastuzumab+pertuzumab+paclitaxel</td>
<td>13 (34.21)</td>
<td>10 (18.87)</td>
<td>23 (25.27)</td>
<td></td>
</tr>
<tr>
<td>Ado-trastuzumab emtansine</td>
<td>3 (7.89)</td>
<td>0 (0)</td>
<td>3 (3.3)</td>
<td></td>
</tr>
<tr>
<td><strong>ECOG PSd, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.26c</td>
</tr>
<tr>
<td>0</td>
<td>13 (34.21)</td>
<td>16 (30.77)b</td>
<td>29 (32.22)b</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15 (39.47)</td>
<td>29 (55.77)b</td>
<td>44 (48.89)b</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6 (15.79)</td>
<td>4 (7.69)b</td>
<td>10 (11.11)b</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2 (5.26)</td>
<td>3 (5.77)b</td>
<td>5 (5.56)b</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2 (5.26)</td>
<td>0 (0)b</td>
<td>2 (2.22)b</td>
<td></td>
</tr>
</tbody>
</table>

**a**Student t test.  
**b**Data missing for 1 participant.  
**c**χ² test.  
**d**ECOG PS: Eastern Cooperative Oncology Group Performance Status.

**ePRO Data**

In the prospective cohort, 84 of the 92 available different symptoms were entered (average >4 symptoms/day), resulting in a total of 9680 symptoms, whereas 54 of the 82 different symptoms were reported in the historical cohort (average >3 symptoms/day), resulting in a total of 6904 symptom entries. The most common symptoms reported in both groups included fatigue, taste disorder, nausea, diarrhea, dry mucosa, joint discomfort, tingling, sleep disorder, headache, and appetite loss (Figure 1).
Overall, the distribution of symptom grades in the Ogivri cohort revealed that most patients experienced minimal (grade 0) and mild (grade 1) toxicities, followed by grade 2, grade 3, and grade 4 toxicities (Table 2). The results for QoL (based on the EQ-5D-5L questionnaire), which was also assessed in this study, will be reported elsewhere.

Table 2. Distribution of symptom grades in the Ogivri prospective cohort (N=9680 symptoms reported).

<table>
<thead>
<tr>
<th>App symptom score and grade</th>
<th>Entries, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤2 (Grade 0=minimal)</td>
<td>4167 (43.1)</td>
</tr>
<tr>
<td>&gt;2 to ≤4 (Grade 1=mild)</td>
<td>4040 (41.7)</td>
</tr>
<tr>
<td>&gt;4 to ≤6 (Grade 2=moderate)</td>
<td>1268 (13.1)</td>
</tr>
<tr>
<td>&gt;6 to ≤8 (Grade 3=severe)</td>
<td>164 (1.7)</td>
</tr>
<tr>
<td>&gt;8 to ≤10 (Grade 4=very severe)</td>
<td>41 (0.4)</td>
</tr>
</tbody>
</table>

CTCAE Score

The primary endpoint was analyzed in 50 patients (3 patients were excluded due to missing data on symptoms) treated with Ogivri (prospective cohort) and in all 38 patients treated with Herceptin (historical cohort). The mean CTCAE scores were comparable between the two cohorts (Table 3) with a mean difference of –1.27 (95% CI –7.24 to 4.70; P=.68) (Figure 2). In the multivariate analysis, the adjusted mean CTCAE scores also did not differ between the two cohorts (2.51, 95% CI –3.27 to 8.29) (Table S1 in Multimedia Appendix 1).
Figure 2. CTCAE score in the prospective (Ogivri) and historical (Herceptin) cohorts. The CTCAE score (primary endpoint) was analyzed in 50 patients (3 patients were excluded due to missing data on symptoms) treated with Ogivri and in all 38 patients treated with Herceptin. CTCAE: Common Terminology Criteria for Adverse Events.

Table 3. Descriptive statistics of treatment side effects according to Common Terminology Criteria for Adverse Events (CTCAE) scores and well-being according to the Eastern Cooperative Oncology Group Performance Status (ECOG PS).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Historical cohort (n=38)</th>
<th>Prospective cohort (n=53)</th>
<th>Total (N=91)</th>
<th>( P ) value(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTCAE score(^b)</td>
<td></td>
<td></td>
<td></td>
<td>.68</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>30.29 (11.618)</td>
<td>29.02 (15.804)</td>
<td>29.57 (14.088)</td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>32 (1-61)</td>
<td>28 (1-100)</td>
<td>29.5 (1-100)</td>
<td></td>
</tr>
<tr>
<td>ECOG PS(^c)</td>
<td></td>
<td></td>
<td></td>
<td>.28</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>69.82 (23.006)</td>
<td>74.27 (15.66)</td>
<td>72.39 (19.117)</td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>76 (0-100)</td>
<td>74.5 (35-100)</td>
<td>76 (0-100)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Reported \( P \) values correspond to mean differences between cohorts.

\(^b\)Missing scores for 3 participants in the prospective cohort.

\(^c\)Missing score for 1 participant in the prospective cohort.

Well-Being Score

The secondary endpoint, the well-being score, was analyzed in 52 patients (one patient was excluded due to missing data on well-being) from the OGIPRO study and in all 38 patients from the historical cohort. The mean well-being score did not differ significantly between patients treated with Ogivri and those treated with Herceptin (Table 3), with a mean difference of 4.45 (95% CI –3.53 to 12.44; \( P= .28 \)). The adjusted mean well-being scores also did not differ between the two cohorts (3.78, 95% CI –4.64 to 12.19) (Table S2 in Multimedia Appendix 1).

Cognitive Abilities in the Prospective Cohort

A total of 767 cognitive tests were entered and the data of 37 patients (70%) who had performed at least one test were included in the analysis (see Figure S1 in Multimedia Appendix 1). Overall, the mean execution time was 42.9 (SD 26.3) with an absolute difference between the maximum and minimum execution time of 197 seconds. Because of the low sample size and limited number of cognitive tests recorded, no correlation analysis between cognitive abilities and treatment was performed.
**Discussion**

The treatment of patients with HER2-positive BC with the trastuzumab biosimilar Ogivri resulted in equivalent symptoms, AEs, and well-being to those experienced under treatment with Herceptin as determined by ePROs. Ogivri treatment in the context of HER2-positive BC was well tolerated and no new important safety risks were observed. The results of this study are consistent with previously reported evidence on the safety comparability of the trastuzumab biosimilar Ogivri to the reference product Herceptin for the treatment of HER2-positive BC [6,17].

The use of biosimilars in oncology could reduce health care costs and thus expand access to drugs worldwide. The European Medicines Agency as well as the US Food and Drug Administration have developed guidelines requiring biosimilars to demonstrate comparable results in relevant clinical trials to those obtained using the original product [18]. Recent studies have demonstrated that anti-HER2 therapy can be switched safely to trastuzumab biosimilars and successfully implemented in clinical practice [19].

In our study, the incidence and distribution of symptoms associated with Ogivri were similar to those reported with Herceptin. However, the slightly lower mean symptom score related to Ogivri might be attributed to the higher number of treatments in this cohort for advanced cancer stages, including antihormone treatments and dual HER2 blockade.

To our knowledge, this study represents the first real-world evaluation on efficacy and safety in patients treated with HER2 biosimilars using ePRO data. Use of the app in this study was intended to help patients gain a better overview of their disease history and improve their symptom management. Our analysis of ePRO data demonstrated comparable CTCAE scores between the prospective Ogivri cohort and the historical Herceptin cohort. These findings further support the previously reported similar safety profiles between the trastuzumab biosimilar and the corresponding reference product [6,17] with no new safety concerns observed.

Importantly, the well-being score based on the ECOG PS did not differ between the two cohorts. In a pooled analysis of data from three randomized clinical trials including patients with HER2-positive advanced BC, PROs were identified as an independent prognostic factor for both survival and toxicity outcomes. In addition, patient-reported physical well-being and clinically interpreted ECOG PS provided independent prognostic information [20]. In our prospective Ogivri cohort, we did not focus on the prognostic value of the ePRO with regard to clinical outcomes, but we were able to demonstrate that an eHealth patient empowerment app can provide reliable information on side effects and well-being when comparing a biosimilar with reference treatments. Hence, the use of continuous eHealth-based symptom reporting together with biosimilars can result in a potential economic benefit by reducing the cost of drug treatment and hospitalization. Further detailed analyses of randomized trials with biosimilars will help to quantify these resources more comprehensively.

In general, the diary characteristic of apps might appear helpful to capture and recall disease-related information such as cognitive impairments [21]. In the OGIPO study, patients had the possibility to complete a TMT, which is one of the most widely used neuropsychological tests in clinical practice: this test is perceptive, easy to understand for patients, has a short administration time, and has shown consistent results in multiple clinical populations [22-24]. A study investigating the impact of chemotherapy on cognitive functions of patients with BC demonstrated increased cognitive impairment throughout chemotherapy treatment, which did not recover 2 months after chemotherapy was completed [25]. In contrast, in the OGIPO study, the cognitive performance of the patients receiving Ogivri showed potential improvement throughout the study treatment. However, due to the low number of cognitive tests recorded during app use, the cognitive abilities were analyzed descriptively and no association could be made with regard to the trastuzumab biosimilar treatment. Further analyses are needed to evaluate the electronically collected cognitive test results in patients treated with biosimilars and corresponding reference products.

Our study has several strengths and limitations. The limitations of the study included the design that lacked a prospective control group so that the study was not randomized. However, the comparison between prospectively collected data of patients treated with Ogivri and the historical ePRO data of patients treated with Herceptin in two previous studies [7,13] demonstrated no difference with regard to symptoms, well-being, and AEs. The earlier versions of the mobile app used in the historical cohort were developed to record symptoms and treatment side effects continuously and according to the CTCAE in patients with cancer, but were not designed to send questionnaires to patients. Nevertheless, the ePRO data of the historical cohort were recorded in the same way in the earlier versions of the mobile app [11] and are thus comparable to those of the prospective cohort. An exploratory analysis on cognitive abilities was performed only in the prospective cohort as these data were not available in the historical cohort. Further studies that are randomized and sufficiently powered to evaluate the real-word cognitive functions in patients with HER2-positive BC treated with anti-HER2 biosimilars are needed.

The major strength of our proof-of-concept study is that it was able to provide the first evidence that data collected via an autonomous eHealth app can also be used longitudinally to determine the similarity of a trastuzumab biosimilar to the reference product for the treatment of HER2-positive BC. Furthermore, our study has reached its primary endpoint, showing a similar average CTCAE score between patients treated with the trastuzumab biosimilar Ogivri and those treated with the reference drug Herceptin. Our results suggest that the use of a patient empowerment eHealth app in patients treated with anti-HER2 biosimilars is reliable and can support therapy management.

In conclusion, in patients with HER2-positive BC, treatments with the trastuzumab biosimilar Ogivri and the reference drug Herceptin resulted in equivalent symptoms, AEs, and well-being reported by ePRO. Hence, the integration of an ePRO tool into research and clinical practice can provide reliable information.
when investigating the real-world tolerability and safety outcomes of similar therapeutic compounds.

**Acknowledgments**

The authors thank all patients who participated in this study along with the investigators and their teams. We also thank Palleos Healthcare for the continued support of the trial; Dr. Stefanie von Felten at University of Zurich, Epidemiology, Biostatistics and Prevention Institute for the assistance with data analysis; and Swiss Tumor Institute, Zurich, Switzerland for the financial support for the trial.

**Data Availability**

The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

**Conflicts of Interest**

AT received medical writing support from Palleos Healthcare, funding from the Swiss Tumor Institute, payment or honoraria for presentations from Viatris, support for attending ESMO 2023 from Viatris, and is the founder and stock owner of Mobile Health AG. YK reports stock or stock options from Viatris and is the Head of Project Management at Mobile Health AG. GAKU reports stock or stock options from Novartis. MA reports consulting fees from Aptar. AE reports consulting fees from Daiichi-Sankyo, Gilead, Merck, Novartis, and Seagen, and institutional financial support for clinical trials from AstraZeneca, Roche, Pfizer, and Novartis. All other authors have declared no conflicts of interest.

**Multimedia Appendix 1**

Multivariate analyses of Common Terminology Criteria for Adverse Events (CTCAE) scores (Table S1) and well-being scores (Table S2); distribution of cognitive performance scores (Figure S1).

**References**


Abbreviations

AE: adverse event  
BC: breast cancer  
BW: body weight  
CTCAE: Common Terminology Criteria for Adverse Events  
ECOG PS: Eastern Cooperative Oncology Group Performance Status  
ePRO: electronic patient-reported outcome  
HER2: human epidermal growth factor 2  
QoL: quality of life  
SAP: statistical analysis plan  
TMT: Trail Making Test
Variation in Trust in Cancer Information Sources by Perceptions of Social Media Health Mis- and Disinformation and by Race and Ethnicity Among Adults in the United States: Cross-Sectional Study

Jim P Stimpson1, PhD; Sungchul Park2, PhD; Sandi L Pruitt1, PhD; Alexander N Ortega3, PhD

1Peter O’Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas, TX, United States
2Department of Health Policy and Management, Korea University, Seoul, Republic of Korea
3Thompson School of Social Work and Public Health, University of Hawai‘i at Mānoa, Honolulu, HI, United States

Corresponding Author:
Jim P Stimpson, PhD
Peter O’Donnell Jr. School of Public Health
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd
Dallas, TX, 75390
United States
Phone: 1 (214) 645 2567
Email: james.stimpson@utsouthwestern.edu

Abstract

Background: Mis- and disinformation on social media have become widespread, which can lead to a lack of trust in health information sources and, in turn, lead to negative health outcomes. Moreover, the effect of mis- and disinformation on trust in information sources may vary by racial and ethnic minoritized populations.

Objective: We evaluated how trust in multiple sources of cancer information varied by perceptions of health mis- and disinformation on social media and by race and ethnicity.

Methods: Cross-sectional, nationally representative survey data from noninstitutionalized adults in the United States from the 2022 Health Information National Trends Survey 6 (HINTS 6) were analyzed (N=4137). The dependent variable measured the level of trust in cancer information sources. The independent variables were perceptions about health mis- and disinformation on social media and race and ethnicity. Multivariable logistic regression models were adjusted for survey weight and design, age, birth gender, race and ethnicity, marital status, urban/rural designation, education, employment status, feelings about household income, frequency of social media visits, and personal and family history of cancer. We also tested the interaction effect between perceptions of social media health mis- and disinformation and participants’ self-reported race and ethnicity.

Results: Perception of “a lot of” health mis- and disinformation on social media, relative to perception of “less than a lot,” was associated with a lower likelihood of high levels of trusting cancer information from government health agencies (odds ratio [OR] 0.60, 95% CI 0.47-0.77), family or friends (OR 0.56, 95% CI 0.44-0.71), charitable organizations (OR 0.78, 95% CI 0.63-0.96), and religious organizations and leaders (OR 0.64, 95% CI 0.52-0.79). Among White participants, those who perceived a lot of health mis- and disinformation on social media were less likely to have high trust in cancer information from government health agencies (margin=61%, 95% CI 57%-66%) and family or friends (margin=49%, 95% CI 43%-55%) compared to those who perceived less than a lot of health mis- and disinformation on social media. Among Black participants, those who perceived a lot of health mis- and disinformation on social media were less likely to have high trust in cancer information from religious organizations and leaders (margin=20%, 95% CI 10%-30%) compared to participants who perceived no or a little health mis- and disinformation on social media.

Conclusions: Certain sources of cancer information may need enhanced support against the threat of mis- and disinformation, such as government health agencies, charitable organizations, religious organizations and leaders, and family or friends. Moreover, interventions should partner with racial and ethnically minoritized populations that are more likely to have low trust in certain cancer information sources associated with mis- and disinformation on social media.
Introduction

Misinformation is unintentionally providing false or inaccurate information, while disinformation is intentionally spreading false or inaccurate information [1-3]. A recent systematic review found that more than 80% of adult social media users perceive “some” or “a lot of” false or misleading health information on social media, while nearly a fifth reported either “none” or “a little” [4]. Both mis- and disinformation have been linked to reductions in health-promoting behaviors. For example, people who perceive more misinformation in the media are associated with a lower likelihood of being vaccinated against COVID-19 and a greater likelihood of smoking more and having poorer nutrition than people who perceive less misinformation in the media [5-11]. According to the Comprehensive Model of Information Seeking, misinformation may be associated with a lack of trust in health information sources, which can, in turn, lead to changes in health behaviors [12,13].

There is limited research on misinformation and trust, with some mixed findings. Some cross-sectional studies have found that higher perceptions of misinformation are associated with lower trust in the media, while one study of multiple countries, including the United States, did not find a relationship between perceptions of misinformation and trust in news media [14-19]. A gap in the literature is that these studies were not drawn from representative samples and only measured trust in media. Furthermore, the effects of misinformation may be more pronounced among individuals with comorbidities, particularly cancer, that have complex clinical treatment plans and significant economic costs [20]. For example, cancer survivors are more likely to have a lot of trust in information from doctors compared to persons that have not been diagnosed with cancer [20]. Therefore, there is an evidence gap for the effects of social media mis- and disinformation on trust in different credible sources (eg, scientists, doctors, and government health agencies) of cancer information.

The effect of mis- and disinformation on trust may also vary by different population groups. In some studies, racial and ethnic minoritized populations were found to be less likely to perceive false or misleading health information on social media and to trust noncredible information sources compared to non-Latino White people [21,22]. The lack of trust may also extend to credible sources of cancer information because, for example, non-Latino Black and Latino people have reported lower trust in doctors compared to non-Latino White people [20,23]. A study of 10-year trends in trust in cancer information found that, compared to non-Latino White participants, non-Latino Black participants were more likely to trust cancer information from government, charitable organizations, and religious organizations. In contrast, that same study found that Latino participants were less likely to trust cancer information from doctors compared to non-Latino White people [24]. There may be differences within Latino populations in trust in cancer information. For example, Cuban Americans and Puerto Ricans were more than twice as likely to trust information about cancer from print media and religious organizations compared to Mexican Americans [25]. However, a recent study found that trust in cancer information from government health agencies and family or friends declined among non-Latino Black participants from 2018 to 2020 [26]. Given these mixed findings, there is a need to examine whether the effect of mis- and disinformation on trust in cancer information varies among racial and ethnic minoritized populations and therefore may be a mechanism to explain these variations and a possible target for interventions to improve trust in cancer information, at least from credible sources such as doctors and scientists [27].

Research Objective

The purpose of this study is to use recently released nationally representative data to estimate the association between perceptions of false or misleading health information on social media and level of trust in multiple sources of information about cancer. We hypothesized that perception of a lot of health mis- and disinformation on social media would be associated with lower levels of trust in cancer information sources. By extension, this study evaluated the interaction effect between race and ethnicity of the participants, perceptions of social media health mis- and disinformation, and trust in cancer information. We hypothesized that the association between perceptions of a lot of mis- and disinformation on social media and trust in cancer information sources would vary by race and ethnicity. The results of this research have implications for effective communication about cancer in public health education campaigns, especially for racial and ethnic minoritized populations.

Methods

Data

This study used cross-sectional data from the Health Information National Trends Survey 6 (HINTS 6), which is a nationally representative survey of civilian, noninstitutionalized adults aged 18 years and older living in the United States. HINTS 6 provides data on adults’ knowledge of cancer risk factors, attitudes toward cancer screening, and cancer prevention and screening behaviors. HINTS 6 used a 2-stage probability sample of residential addresses. Mail and online surveys were administered to household members from March 7 to November 8, 2022, with a response rate of 28.1% [28]. The data are publicly available and deidentified. Further details about the survey methodology and recruitment procedures are available from the HINTS 6 Methodology Report [28].

Given the focus of this study was perceptions of false or misleading health information on social media, persons that reported that they did not use social media were excluded. There were 4710 cases with complete data for the dependent and
independent variables. After using listwise deletion for 573 cases with missing data for the control variables, the final analytical sample consisted of 4137 adult social media users.

**Measures**

Our dependent variables were measured by asking participants, “In general, how much would you trust information about cancer from...?” Responses included the following: “a doctor,” “family or friends,” “religious organizations and leaders,” “government health agencies,” “charitable organizations,” and “scientists.” The response options were dichotomized into low levels of trust (“not at all” or “a little”) versus high levels of trust (“some” or “a lot”).

The primary independent variable was perceptions about health mis- and disinformation on social media, which was assessed by the following question: “How much of the health information that you see on social media do you think is false or misleading?” HINTS had not measured perceptions about social media mis- and disinformation in prior iterations of the survey. However, this measure did not differentiate between people’s perceptions of mis- versus disinformation. The original response categories were “a lot,” “some,” “a little,” and “none.” We dichotomized this as “less than a lot” (including “some,” “a little,” and “none”) versus “a lot.” Race and ethnicity were self-reported by the participants in 5 categories: “non-Latino White,” “non-Latino Black,” “Asian American,” “other,” and “Latino.”

Demographic control variables included age (18-34, 35-49, 50-64, and ≥65 years), sex (male and female), marital status (married or cohabiting, formerly married, and never married), residence in a metropolitan versus nonmetropolitan county as designated by the United States Department of Agriculture in 2013, education (high school or less, some college, and college degree or higher), full-time employment status, and feelings about household income (finding it very difficult on present income, getting by on present income, and living comfortably on present income). It should be noted that age was not collected as a continuous variable in HINTS 6, which limited the age categories that could be analyzed. In addition, we controlled for frequency of visiting social media sites (never, monthly/weekly, and daily) in the past 12 months and personal and family (first- or second-degree biological relatives) history of cancer.

**Statistical Analysis**

All analyses accounted for survey weights and design using jackknife replicate weights for variance estimation. Statistical significance was set at \( \alpha < .05 \). The descriptive statistics for the study sample were calculated as survey-weighted percentages accompanied with the raw sample size for each variable. The bivariate relationship between level of trust in cancer information and perceptions of mis- and disinformation were calculated with cell percentages and adjusted Wald \( P \) values. Then, multivariable logistic regression models were calculated for each dichotomous outcome. In addition to the main effect, we also tested the interaction effect between perceptions of health mis- and disinformation on social media and participants’ self-reported race and ethnicity. To facilitate interpretation of the interaction effect, we calculated predicted marginal effects from the multivariable logistic regression models.

For this study, the primary focus was perceptions of information on social media. Therefore, we conducted a sensitivity analysis in which we excluded 257 adults who had not visited a social media site in the past year or reported that they did not use social media (n=3880 were included). After excluding these participants, the results were similar, as shown in Multimedia Appendix 1, Table S1. In addition, we conducted a sensitivity analysis for an ordinal measurement of the dependent variables (“a lot,” “some,” “a little,” and “not at all”) using ordered logit regression, and we found that the results were replicated with this alternative measurement, as shown in Multimedia Appendix 1, Table S2. Another sensitivity analysis included participants that did not use social media (n=4986). After including participants that did not use social media, the results were similar, as shown in Multimedia Appendix 1, Table S3. In Multimedia Appendix 1, Table S4, we tested an alternative measurement of the independent variable in which perception of “a lot” of social media mis- and disinformation was compared with respondents that reported “some” and “none” or “a little.” For this sensitivity analysis, we combined “none” and “a little” because only 108 participants chose “none” for this measure. We replicated the main result using this alternative measurement of the independent variable.

**Ethical Considerations**

The University of Texas Southwestern Medical Center institutional review board determined that the study was exempt from review because it used publicly available data without personal identifiers.

**Results**

Table 1 provides the survey-weighted percentages for the study variables. Most participants in the survey reported high trust in cancer information from doctors (95%), scientists (86%), and government health agencies (71%). About half reported high trust in cancer information from family or friends (54%) and charitable organizations (49%). About a quarter of participants reported high trust in cancer information from religious organizations and leaders (26%). When participants were asked about perceptions of false or misleading health information on social media, most reported “less than a lot” (63%) and 37% reported “a lot.”
Table 1. Unadjusted sample size and survey-weighted percentages for study variables from the 2022 Health Information National Trends Survey 6 (N=4137).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unadjusted sample size, n (weighted %)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome variables</strong></td>
<td></td>
</tr>
<tr>
<td>In general, how much would you trust information about cancer from a doctor?</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>200 (5)</td>
</tr>
<tr>
<td>High</td>
<td>3937 (95)</td>
</tr>
<tr>
<td>In general, how much would you trust information about cancer from scientists?</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>525 (14)</td>
</tr>
<tr>
<td>High</td>
<td>3612 (86)</td>
</tr>
<tr>
<td>In general, how much would you trust information about cancer from government health agencies?</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1077 (29)</td>
</tr>
<tr>
<td>High</td>
<td>3060 (71)</td>
</tr>
<tr>
<td>In general, how much would you trust information about cancer from family or friends?</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1874 (46)</td>
</tr>
<tr>
<td>High</td>
<td>2263 (54)</td>
</tr>
<tr>
<td>In general, how much would you trust information about cancer from charitable organizations?</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2050 (51)</td>
</tr>
<tr>
<td>High</td>
<td>2087 (49)</td>
</tr>
<tr>
<td>In general, how much would you trust information about cancer from religious organizations and leaders?</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>3034 (74)</td>
</tr>
<tr>
<td>High</td>
<td>1103 (26)</td>
</tr>
<tr>
<td><strong>Independent variables</strong></td>
<td></td>
</tr>
<tr>
<td>How much of the health information that you see on social media do you think is false or misleading?</td>
<td></td>
</tr>
<tr>
<td>Less than a lot</td>
<td>2643 (63)</td>
</tr>
<tr>
<td>A lot</td>
<td>1494 (37)</td>
</tr>
<tr>
<td><strong>Race and ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>Non-Latino White</td>
<td>2381 (61)</td>
</tr>
<tr>
<td>Non-Latino Black</td>
<td>643 (11)</td>
</tr>
<tr>
<td>Latino</td>
<td>734 (17)</td>
</tr>
<tr>
<td>Non-Latino Asian American</td>
<td>230 (6)</td>
</tr>
<tr>
<td>Non-Latino other</td>
<td>149 (5)</td>
</tr>
<tr>
<td><strong>Age group (years)</strong></td>
<td></td>
</tr>
<tr>
<td>18-34</td>
<td>771 (29)</td>
</tr>
<tr>
<td>35-49</td>
<td>1012 (29)</td>
</tr>
<tr>
<td>50-64</td>
<td>1222 (27)</td>
</tr>
<tr>
<td>≥65</td>
<td>1132 (15)</td>
</tr>
<tr>
<td><strong>Birth gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1586 (48)</td>
</tr>
<tr>
<td>Female</td>
<td>2551 (52)</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>2290 (57)</td>
</tr>
<tr>
<td>Formerly married</td>
<td>994 (10)</td>
</tr>
<tr>
<td>Never married</td>
<td>853 (33)</td>
</tr>
<tr>
<td>Variables</td>
<td>Unadjusted sample size, n (weighted %)</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td><strong>United States Department of Agriculture 2013 rural/urban designation</strong></td>
<td></td>
</tr>
<tr>
<td>Nonmetropolitan</td>
<td>512 (12)</td>
</tr>
<tr>
<td>Metropolitan</td>
<td>3625 (88)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>High school or less</td>
<td>812 (25)</td>
</tr>
<tr>
<td>Some college</td>
<td>1185 (39)</td>
</tr>
<tr>
<td>College graduate or higher</td>
<td>2140 (36)</td>
</tr>
<tr>
<td><strong>Work full time (past 30 days)</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1878 (40)</td>
</tr>
<tr>
<td>Yes</td>
<td>2259 (60)</td>
</tr>
<tr>
<td><strong>Feelings about household income</strong></td>
<td></td>
</tr>
<tr>
<td>Finding it very difficult on present income</td>
<td>811 (19)</td>
</tr>
<tr>
<td>Getting by on present income</td>
<td>1505 (36)</td>
</tr>
<tr>
<td>Living comfortably on present income</td>
<td>1821 (45)</td>
</tr>
<tr>
<td><strong>Frequency of social media site visits</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>257 (6)</td>
</tr>
<tr>
<td>Monthly or weekly</td>
<td>1119 (25)</td>
</tr>
<tr>
<td>Daily</td>
<td>2761 (70)</td>
</tr>
<tr>
<td><strong>Personal history of cancer</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>3593 (91)</td>
</tr>
<tr>
<td>Yes</td>
<td>544 (9)</td>
</tr>
<tr>
<td><strong>Family history of cancer</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1259 (35)</td>
</tr>
<tr>
<td>Yes</td>
<td>2878 (65)</td>
</tr>
</tbody>
</table>

Table 2 provides the bivariable relationship between the outcome variables and the independent variable. There was not a statistically significant relationship between perception of mis- and disinformation and trust in cancer information from doctors ($P=.93$) or scientists ($P=.85$). However, there was a statistically significant bivariable relationship between perception of mis- and disinformation and trust in cancer information from government health agencies ($P<.001$), family or friends ($P<.001$), charitable organizations ($P=.007$), and religious organizations and leaders ($P<.001$). About a quarter of participants (24%) that perceived a lot of mis- and disinformation on social media had a high level of trust in government health agencies. Nearly half of participants (47%) that perceived less than a lot of mis- and disinformation on social media had a high level of trust in government health agencies. Only 17% of participants that perceived a lot of mis- and disinformation on social media had a high level of trust in family or friends. In contrast, 37% of participants that perceived less than a lot of mis- and disinformation on social media had a high level of trust in family or friends. Only 16% of participants that perceived a lot of mis- and disinformation on social media had a high level of trust in charitable organizations. A third of participants (33%) that perceived less than a lot of mis- and disinformation on social media had a high level of trust in charitable organizations. Finally, only 7% of participants that perceived a lot of mis- and disinformation on social media had a high level of trust in charitable organizations. Finally, only 7% of participants that perceived a lot of mis- and disinformation on social media had a high level of trust in charitable organizations. Nearly 1 in 5 participants (19%) that perceived less than a lot of mis- and disinformation on social media had a high level of trust in religious organizations and leaders.
Table 2. Survey-weighted unadjusted bivariable relationship between trust in cancer information source (low vs high) and perception of health mis- and disinformation on social media ("less than a lot" vs "a lot") from the 2022 Health Information National Trends Survey 6 (N=4137).

<table>
<thead>
<tr>
<th>Cancer information source</th>
<th>Trust in cancer information source and perception of health mis- and disinformation on social media</th>
<th>( P ) value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low trust</td>
<td>High trust</td>
</tr>
<tr>
<td></td>
<td>Less than a lot&lt;sup&gt;b&lt;/sup&gt;, %</td>
<td>A lot&lt;sup&gt;c&lt;/sup&gt;, %</td>
</tr>
<tr>
<td>Doctors</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>.93</td>
<td></td>
</tr>
<tr>
<td>Scientists</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>.85</td>
<td></td>
</tr>
<tr>
<td>Government health agencies</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Family or friends</td>
<td>26</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Charitable organizations</td>
<td>31</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>Religious organizations and leaders</td>
<td>45</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>P values were calculated with the adjusted Wald \( \chi^2 \) test.

<sup>b</sup>Perception of “less than a lot” of health mis- and disinformation on social media.

<sup>c</sup>Perception of “a lot” of health mis- and disinformation on social media.

Table 3 provides the multivariable odds ratios (ORs) and 95% CIs calculated from logistic regression. Perception of a lot of health mis- and disinformation on social media, relative to perception of less than a lot, was associated with a lower likelihood of high levels of trusting cancer information from government health agencies (OR 0.60, 95% CI 0.47-0.77), family or friends (OR 0.56, 95% CI 0.44-0.71), charitable organizations (OR 0.78, 95% CI 0.63-0.96), and religious organizations and leaders (OR 0.64, 95% CI 0.52-0.79). There was not a statistically significant association between perception of social media health mis- and disinformation and level of trust in cancer information from doctors (OR 0.95, 95% CI 0.45-2.01) or scientists (OR 0.98, 95% CI 0.72-1.33).

Table 3. Multivariable odds ratios (ORs) and 95% CIs for perceptions of social media health mis- and disinformation and trust in cancer information sources from the 2022 Health Information National Trends Survey 6 (N=4137). Logistic regression models were adjusted for survey weight and design, age, birth gender, marital status, urban or rural designation, race and ethnicity, education, employment status, feelings about household income, frequency of social media visits, and personal and family history of cancer.

<table>
<thead>
<tr>
<th>Cancer information source</th>
<th>Trust in cancer information source among participants with the perception that a lot of health information on social media is false or misleading, odds ratio&lt;sup&gt;a&lt;/sup&gt; (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctors</td>
<td>0.95 (0.45-2.01)</td>
</tr>
<tr>
<td>Scientists</td>
<td>0.98 (0.72-1.33)</td>
</tr>
<tr>
<td>Government health agencies</td>
<td>0.60 (0.47-0.77)</td>
</tr>
<tr>
<td>Family or friends</td>
<td>0.56 (0.44-0.71)</td>
</tr>
<tr>
<td>Charitable organizations</td>
<td>0.78 (0.63-0.96)</td>
</tr>
<tr>
<td>Religious organizations and leaders</td>
<td>0.64 (0.52-0.79)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Reference: “less than a lot.”

Table 4 provides the predicted marginal effects, interpreted as percentage points, calculated from the multivariable logistic regression–adjusted interaction effects between perceptions of health mis- and disinformation on social media and participants’ self-reported race and ethnicity. There was not a statistically significant interaction effect between perception of mis- and disinformation, race and ethnicity, and trust in cancer information from doctors or scientists. Among White participants, those who perceived a lot of health misinformation and disinformation on social media were less likely to have high trust in cancer information from government health agencies (margin=61%, 95% CI 57%-66%) and family or friends (margin=49%, 95% CI 43%-55%) compared to those who perceived less than a lot of health mis- and disinformation on social media. Among Black participants, those who perceived a lot of health mis- and disinformation on social media were less likely to have high trust in cancer information from religious organizations and leaders (margin=20%, 95% CI 10%-30%) compared to participants who perceived less than a lot of health mis- and disinformation on social media.

Table 4. The predicted marginal effects, interpreted as percentage points, calculated from the multivariable logistic regression–adjusted interaction effects between perceptions of health mis- and disinformation on social media and participants’ self-reported race and ethnicity.
Table 4. Multivariable-adjusted percentage points for trusting cancer information by source and the interaction effect between race and ethnicity and perceptions of health mis- and disinformation on social media from the 2022 Health Information National Trends Survey 6 (N=4137). Predicted marginal effects were calculated from multivariable logistic regression models that were adjusted for survey weight and design, age, birth gender, marital status, urban or rural designation, education, employment status, feelings about household income, frequency of social media visits, and personal and family history of cancer.

<table>
<thead>
<tr>
<th>Race and ethnicity</th>
<th>Perception of false or misleading health information from cancer information source, percentage points (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Doctors</td>
</tr>
<tr>
<td></td>
<td>Scientists</td>
</tr>
<tr>
<td></td>
<td>Government health agencies</td>
</tr>
<tr>
<td></td>
<td>Family or friends</td>
</tr>
<tr>
<td></td>
<td>Charitable organizations and leaders</td>
</tr>
<tr>
<td>Non-Latino White</td>
<td></td>
</tr>
<tr>
<td>Less than a lot</td>
<td>96 (93-98)</td>
</tr>
<tr>
<td>A lot</td>
<td>95 (93-97)</td>
</tr>
<tr>
<td>Non-Latino Black</td>
<td></td>
</tr>
<tr>
<td>Less than a lot</td>
<td>96 (93-99)</td>
</tr>
<tr>
<td>A lot</td>
<td>96 (92-101)</td>
</tr>
<tr>
<td>Latino</td>
<td></td>
</tr>
<tr>
<td>Less than a lot</td>
<td>93 (89-98)</td>
</tr>
<tr>
<td>A lot</td>
<td>96 (92-100)</td>
</tr>
<tr>
<td>Non-Latino Asian American</td>
<td></td>
</tr>
<tr>
<td>Less than a lot</td>
<td>89 (61-117)</td>
</tr>
<tr>
<td>A lot</td>
<td>99 (93-105)</td>
</tr>
<tr>
<td>Non-Latino Other</td>
<td></td>
</tr>
<tr>
<td>Less than a lot</td>
<td>96 (84-107)</td>
</tr>
<tr>
<td>A lot</td>
<td>85 (70-99)</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

We found that trust in cancer information from doctors or scientists did not vary based on perceptions of health mis- and disinformation on social media. This suggests that people view doctors and scientists as credible sources of cancer information. However, we found that perception of a lot of mis- and disinformation was associated with reduced levels of trust in cancer information from family or friends, government health agencies, charitable organizations, and religious organizations and leaders. This finding supports other studies that found that mis- and disinformation is associated with reductions in trust in media but extends this prior literature by finding an impact on trust in other sources of cancer information [14-19]. Moreover, this finding is consistent with the Comprehensive Model of Information Seeking, which identifies trust as a mechanism linking mis- and disinformation to health behaviors [12,13].

There were notable variations in the relationship between trust in cancer information sources, perceptions of false or misleading health information, and race and ethnicity. For instance, we found that Black participants who perceived a lot of health mis- and disinformation on social media were less likely to have high trust in cancer information from religious organizations and leaders compared to Black participants who perceived less than a lot of health mis- and disinformation on social media. Another contribution of our study is that White participants who perceived a lot of health mis- and disinformation on social media were less likely to have high trust in cancer information from government health agencies and family or friends compared to White participants who perceived less than a lot of health mis- and disinformation on social media. There have been mixed findings on trust in cancer information sources by race and ethnicity in the recent literature, with one study finding higher trust among Black participants for several sources of cancer information compared to White participants and lower trust in doctors among Latino participants compared to White participants [21-25]. However, another study found that trust in cancer information from government health agencies and family or friends declined among Black participants after the COVID-19 pandemic [26]. Our study adds to this literature by identifying that the effect of mis- and disinformation on trusting information sources may vary among racial and ethnic minoritized populations.

Limitations

We were able to replicate the findings of the study using several different sensitivity analyses, as shown in Multimedia Appendix 1. However, the results should be interpreted within the constraints of the cross-sectional data. First, this study cannot be used to determine the causal relationship between perceptions of mis- and disinformation and trust in social institutions. Second, the 2022 wave of the HINTS survey was the first time that the public’s perceptions of mis- and disinformation were
measured. If this measure is collected in subsequent iterations of HINTS, then analyses may be able to detect changes in the association between mis- and disinformation and trust in information sources over time. We note that perceptions of mis- and disinformation may not be an accurate measure of objective exposure to social media mis- and disinformation. Further, this measure does not differentiate between people’s perceptions of mis- versus disinformation. Another limitation is that the focus of this study was on social media mis- and disinformation rather than all media, such as traditional television and print, and therefore the results should be interpreted for this specific form of media. Finally, this study focused on trust in cancer information, and the findings might not apply to trust in other types of health information. By extension, levels of trust in government information may differ between federal and state government health agencies, which were not differentiated in our study [29,30].

**Conclusion**

Certain sources of cancer information may need enhanced support from the threat of mis- and disinformation, such as government health agencies, charitable organizations, religious organizations and leaders, and family or friends. Moreover, there were notable variations in the relationship between trust in cancer information sources (government health agencies, family or friends, and religious organizations and leaders), perceptions of false or misleading health information, and race and ethnicity. One positive finding is that perceptions of mis- and disinformation were not associated with levels of trust in credible sources of cancer information such as doctors or scientists overall or by race and ethnicity. In prior work, researchers have suggested that interventions should be focused on improving trust in science [1]. Although bolstering trust in science or doctors is important, our findings indicate that other sources of cancer information may be more susceptible to the threat of mis- and disinformation. Moreover, interventions should partner with racial and ethnically minoritized populations that are more likely to have low trust in certain cancer information sources associated with mis- and disinformation on social media.

**Acknowledgments**

This research was supported by the National Institute on Minority Health and Health Disparities (NIMHD; award R01MD018727). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIMHD. The funders had no role in study design, data analysis, decision to publish, or preparation of the manuscript.

**Data Availability**

The data sets generated during and/or analyzed during this study are available in the National Cancer Institute repository [31].

**Authors' Contributions**

All authors contributed to the study conception and design. Data analysis was performed by JPS. The first draft of the manuscript was written by JPS. All authors contributed to subsequent drafts of the manuscript. All authors read and approved the final version of the manuscript.

**Conflicts of Interest**

JPS, SP, and ANO have no relevant financial or nonfinancial conflicts of interest to disclose. Unrelated to this work, SLP receives consulting fees from Pfizer and Gilead.

Multimedia Appendix 1
Supplemental analyses.

[DOCX File, 30 KB - cancer_v10i1e54162_app1.docx ]

**References**


14. Ognyanova K, Lazer D, Robertson RE, Wilson C. Misinformation in action: fake news exposure is linked to lower trust in science and belief in misinformation mediate the effects of political orientation on vaccine hesitancy and intention to be vaccinated. Acta Psychol (Amst) 2023 Jul;237:103945 [FREE Full text] [Medline: 37210865]

15. Boulianne S, Humphreit E. Perceived exposure to misinformation and trust in institutions in four countries before and during a pandemic. Int J Comm 2023 Feb 26:17 [FREE Full text]


Abbreviations

HINTS: Health Information National Trends Survey.
OR: odds ratio.
Original Paper


Yue Zheng1,2*, MD; Ailin Zhao3*, MD; Yuqi Yang4*, MBBS; Ladaoan Wang1,2, MD; Yifei Hu4, MBBS; Ren Luo1,2, MD, PhD; Yijun Wu1,2, MD

1Division of Thoracic Tumor Multimodality Treatment, Cancer Center, West China Hospital, Sichuan University, Chengdu, China
2Laboratory of Clinical Cell Therapy, West China Hospital, Sichuan University, Chengdu, China
3Department of Hematology, West China Hospital, Sichuan University, Chengdu, China
4West China School of Medicine, Sichuan University, Chengdu, China
*these authors contributed equally

Corresponding Author:
Yijun Wu, MD
Division of Thoracic Tumor Multimodality Treatment
Cancer Center, West China Hospital
Sichuan University
Guoxue Lane 37
Chengdu, 610041
China
Phone: 86 17888841669
Email: wuyj01029@wchscu.cn

Abstract

Background: Metachronous second primary lung cancer (MSPLC) is not that rare but is seldom studied.

Objective: We aim to compare real-world survival outcomes between different surgery strategies and radiotherapy for MSPLC.

Methods: This retrospective study analyzed data collected from patients with MSPLC between 1988 and 2012 in the Surveillance, Epidemiology, and End Results (SEER) database. Propensity score matching (PSM) analyses and machine learning were performed to compare variables between patients with MSPLC. Survival curves were plotted using the Kaplan-Meier method and were compared using log-rank tests.

Results: A total of 2451 MSPLC patients were categorized into the following treatment groups: 864 (35.3%) received radiotherapy, 759 (31%) underwent surgery, 89 (3.6%) had surgery plus radiotherapy, and 739 (30.2%) had neither treatment. After PSM, 470 pairs each for radiotherapy and surgery were generated. The surgery group had significantly better survival than the radiotherapy group (P < .001) and the untreated group (563 pairs; P < .001). Further analysis revealed that both wedge resection (85 pairs; P = .004) and lobectomy (71 pairs; P = .002) outperformed radiotherapy in overall survival for MSPLC patients. Machine learning models (extreme gradient boosting, random forest classifier, adaptive boosting) demonstrated high predictive performance based on area under the curve (AUC) values. Least absolute shrinkage and selection operator (LASSO) regression analysis identified 9 significant variables impacting cancer-specific survival, emphasizing surgery’s consistent influence across 1 year to 10 years. These variables encompassed age at diagnosis, sex, year of diagnosis, radiotherapy of initial primary lung cancer (IPLC), primary site, histology, surgery, chemotherapy, and radiotherapy of MPSLC. Competing risk analysis highlighted lower mortality for female MPSLC patients (hazard ratio [HR]=0.79, 95% CI 0.71-0.87) and recent IPLC diagnoses (HR=0.79, 95% CI 0.73-0.85), while radiotherapy for IPLC increased mortality (HR=1.31, 95% CI 1.16-1.50). Surgery alone had the lowest cancer-specific mortality (HR=0.83, 95% CI 0.81-0.85), with sublevel resection having the lowest mortality rate among the surgical approaches (HR=0.26, 95% CI 0.21-0.31). The findings provide valuable insights into the factors that influence cumulative cancer-specific mortality.

Conclusions: Surgical resections such as wedge resection and lobectomy confer better survival than radiation therapy for MSPLC, but radiation can be a valid alternative for the treatment of MSPLC.
metachronous second primary lung cancer; radiotherapy; surgical resection; propensity score matching analysis; machine learning

**Introduction**

Lung cancer has become a leading cause of cancer-related deaths worldwide [1]. With the rapid development of screening tools and therapeutic strategies, survival outcomes of lung cancer patients have encouragingly improved, especially for early-stage non-small cell lung cancer (NSCLC), which has a 5-year survival rate as high as 90% [2]. For cancer survivors, longer survival may well lead to a higher probability of developing a second primary cancer. In recent years, metachronous second primary lung cancer (MSPLC) has been commonly observed among survivors with previously treated lung cancer. Thakur et al [3] reported that MSPLC occurred in 2.95% of patients with initial primary lung cancer (IPLC) in the Surveillance, Epidemiology, and End Results (SEER) database. According to the study by Surapaneni et al [4], the risk of developing a second lung cancer is the highest in the first year and continues to be high at 10 years. The surveillance and management of patients with MSPLC have become an urgent issue.

For patients with an initial, early-stage lung cancer, surgical resection remains the most effective treatment. However, there is still a lack of guidelines to assess tumor resectability in patients with MSPLC. Several studies have confirmed the feasibility of surgery for MSPLC [5–9]. Remarkably, patients with MSPLC with previously resected lung cancer may be in poor physical condition and have insufficient lung function reserve, and another surgical procedure may not be appropriate. Thus, an alternative treatment is required for patients with inoperable MSPLC.

Radiation therapy is an effective treatment choice for patients with MSPLC and has fewer complications and impairments. Stereotactic body radiotherapy has recently been reported to have similar survival outcomes as surgery in patients with early-stage lung cancer [10,11]. Previous studies have shown that radiotherapy is a safe and feasible treatment for MSPLC, but whether it can compare with surgery in terms of survival outcomes remains debated [12,13]. Therefore, in this population-based study, the initial step involved conducting propensity score matching (PSM) analyses to compare the survival outcomes of patients who underwent surgical resection with those who received radiotherapy for multiple synchronous primary lung cancers. Furthermore, specific focus was placed on comparing the outcomes of common surgical methods, namely lobectomy and wedge resection, with those of radiotherapy for patients with MSPLC. To enhance the accuracy of the predictions, state-of-the-art machine learning (ML) techniques were used, and multiple algorithms were used to develop robust prediction models.

**Methods**

**Data Source**

Data for all patients diagnosed with MSPLC included in this retrospective study were sourced from the SEER database [14], covering approximately 30% of cancer patients in the United States. Data pertaining to these patients were extracted from 9 cancer registries and augmented with additional treatment information from regions including Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco–Oakland, Seattle–Puget Sound, and Utah. The data set's most recent follow-up information was updated in November 2018. This study aimed to prognosticate the outcomes for patients with MSPLC. In adherence to the established guidelines for the development and reporting of ML predictive models in biomedical research [15], we meticulously maintained precision and clarity throughout our research process.

**Preparation of Data for Model Building**

Patients aged ≥20 years who were diagnosed with MSPLC were identified from the SEER database. We defined MSPLC according to the criteria set by Martini and Melamed [16]. We only included patients with 2 primary lung tumors with a diagnostic interval between the tumors ≥4 years, because it is difficult to distinguish a primary lung tumor from relapse or metastasis when the interval is <4 years [17]. The initial inclusion criteria were as follows: (1) primary sites of the 2 tumors were the lung and bronchus (International Classification of Diseases for Oncology [ICD-O]-3/World Health Organization [WHO] 2008, Third Edition), (2) the time of diagnosis for the IPLC was from January 1988 to December 2012 (to ensure that all enrolled patients had been followed for enough time), and (3) age was ≥20 years. The exclusion criteria included (1) <4 years between the diagnosis of the 2 primary tumors, (2) distant metastasis, (3) histological type of small cell lung cancer for IPLC or MSPLC, and (4) incomplete follow-up information.

We collected the patients' demographic features and clinical characteristics, such as age at diagnosis, sex, race (White, Black, other [American Indian/Alaska Native, Asian/Pacific Islander], and unknown), location relationship of the 2 primary tumors (ipsilateral and contralateral), diagnostic interval, year of diagnosis, SEER cancer stage (localized and regional), histological type (adenocarcinoma, squamous cell carcinoma, and other NSCLC), grade, surgical procedure, chemotherapy, and radiotherapy (beam radiation). Sublevel resection was regarded as an extent of resection that was less than lobectomy. For patients diagnosed with IPLC after 2004, additional clinical information such as TNM (tumor [T], extent of spread to the lymph nodes [N], and presence of metastasis [M]) stage (6th edition of the American Joint Committee on Cancer TNM system) and tumor size were available.
Predictive Models

We used 6 classical ML algorithms, namely extreme gradient boosting (XGB), random forest classifier (RFC), adaptive boosting (ADB), K nearest neighbor (KNN), artificial neural network (ANN), and gradient boosting decision tree (GBDT), to forecast long-term cancer-specific survival (CSS). To select the variables for modeling, the least absolute shrinkage and selection operator (LASSO) regression technique was used. An extensive method was used to determine the optimal combination of variables for each algorithm. The performance and predictive capabilities of over a dozen variables were individually assessed using the models, measured using the area under the receiver operating characteristic curves (AUC of ROCs), and decision curve analysis was conducted. The most effective variables were identified, and additional variables were combined iteratively until the best overall results were obtained. The selection of the optimal modeling approach for each algorithm was determined using 5-fold cross-validation. Furthermore, the contribution of each variable was calculated. Additionally, age-adjusted competing risk regression analysis was conducted using the “cmprsk” package in R to examine the cumulative risk of cancer-specific mortality. This comprehensive approach facilitated a thorough evaluation of the risk factors and outcomes associated with cancer-specific mortality in diverse patient populations.

Statistical Analyses

All statistical analyses were performed using SPSS 27.0 (IBM Corp) and R software version 4.3.1 [18]. SEER*Stat software version 8.4.2 was used to identify the study population from the SEER database. A 2-tailed *P* value <.05 was considered statistically significant. Continuous parameters such as patients’ age and diagnostic interval are expressed as mean (SD) and were compared between the different treatment groups using Mann-Whitney *U* tests. For categorical parameters, proportions were compared using Pearson chi-square tests. To balance the baseline characteristics between the different treatment groups, PSM analyses were used. Survival curves were plotted using the Kaplan-Meier method and compared using log-rank tests.

Ethical Considerations

The data used in this research were extracted from the publicly accessible, anonymized SEER database. Given the nature of the SEER database, which contains deidentified patient information and is widely used for epidemiological and clinical research purposes, our study fell within the category of research that is exempt from formal ethical approval and consent requirements. This exemption is consistent with established institutional and local policies regarding the use of publicly available, deidentified data for research purposes [19].

Results

Demographic Characteristics

According to our inclusion and exclusion criteria, a total of 2451 patients diagnosed with MSPLC were included in this study. All patients’ baseline characteristics are summarized in Table 1. There were 1137 men and 1314 women, with a mean age of 63.5 (SD 9.2) years. White people accounted for 84.1% (2062/2451) of the study population. The mean diagnostic interval between the 2 primary lung tumors was 101.0 (SD 47.6) months. The year of diagnosis of the IPLC ranged from 1988 to 2012. For IPLC, 264 (10.8%) of the 2451 patients did not undergo any surgical procedure, while 2447 underwent surgical resection, including 295 (295/2447, 12%) sublevel resections, 1786 (1786/2447, 72.9%) lobectomies, and 106 (106/2447, 4.3%) pneumonectomies. Additionally, 465 (465/2451, 19%) patients received chemotherapy, and 489 (489/2451, 20%) underwent radiation therapy for IPLC. Based on treatments for MSPLC, patients were divided into the following 4 subgroups: radiotherapy only (864/2451, 35.3%), surgery only (759/2451, 31%), surgery plus radiotherapy (89/2451, 3.6%), and no treatment (739/2451, 30.2%). The median follow-up time after MSPLC diagnosis was 18 (range: 1-273) months. For the entire study population, the 5-year overall survival (OS) was 34.7%.
Table 1. Demographic and clinical characteristics of 2451 patients diagnosed with second primary lung cancer.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>63.5 (9.2)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>2062 (84.1)</td>
</tr>
<tr>
<td>Black</td>
<td>240 (9.8)</td>
</tr>
<tr>
<td>Other</td>
<td>149 (6.1)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1137 (46.4)</td>
</tr>
<tr>
<td>Female</td>
<td>1314 (53.6)</td>
</tr>
<tr>
<td><strong>Relative location, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Ipsilateral</td>
<td>815 (33.3)</td>
</tr>
<tr>
<td>Contralateral</td>
<td>1636 (66.7)</td>
</tr>
<tr>
<td><strong>Diagnosis interval (months), mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td>101.0 (47.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Initial primary lung cancer</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Year of diagnosis, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>1988-1995</td>
<td>763 (31.1)</td>
</tr>
<tr>
<td>1996-2003</td>
<td>919 (37.5)</td>
</tr>
<tr>
<td>2004-2012</td>
<td>769 (31.4)</td>
</tr>
<tr>
<td><strong>SEER stage, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Localized</td>
<td>1538 (62.7)</td>
</tr>
<tr>
<td>Regional</td>
<td>913 (37.3)</td>
</tr>
<tr>
<td><strong>Histology, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>ADC&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1399 (57.1)</td>
</tr>
<tr>
<td>SCC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>690 (28.2)</td>
</tr>
<tr>
<td>Other NSCLC&lt;sup&gt;d&lt;/sup&gt;</td>
<td>362 (14.8)</td>
</tr>
<tr>
<td><strong>Grade, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>277 (11.3)</td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>844 (34.3)</td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>792 (32.3)</td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>115 (4.7)</td>
</tr>
<tr>
<td>Unknown</td>
<td>423 (17.3)</td>
</tr>
<tr>
<td><strong>Surgery, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No surgery</td>
<td>264 (10.8)</td>
</tr>
<tr>
<td>Sublevel resection</td>
<td>295 (12)</td>
</tr>
<tr>
<td>Lobectomy</td>
<td>1786 (72.9)</td>
</tr>
<tr>
<td>Pneumonectomy</td>
<td>106 (4.3)</td>
</tr>
<tr>
<td><strong>Chemotherapy, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>465 (19)</td>
</tr>
<tr>
<td>No/unknown</td>
<td>1986 (81)</td>
</tr>
<tr>
<td><strong>Radiotherapy, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>489 (20)</td>
</tr>
<tr>
<td>No/unknown</td>
<td>1962 (80)</td>
</tr>
</tbody>
</table>
### Second primary lung cancer

**Surgery, n (%)**
- No surgery: 1603 (65.4)
- Wedge resection: 295 (12)
- Segmentectomy: 61 (2.5)
- Other/inseparable sublevel resection: 87 (3.5)
- Lobectomy: 352 (14.4)
- Pneumonectomy: 53 (2.2)

**Chemotherapy, n (%)**
- Yes: 694 (28.3)
- No/Unknown: 1757 (71.7)

**Radiotherapy, n (%)**
- Yes: 953 (38.9)
- No/Unknown: 1498 (61.1)

**Treatment, n (%)**
- Only radiotherapy: 964 (35.3)
- Only surgery: 759 (31.0)
- Surgery + radiotherapy: 89 (3.6)
- None: 739 (30.2)

---

**Radiotherapy Versus Surgery**

Before PSM, the distributions of several baseline characteristics were significantly different between the radiotherapy and surgery groups. These included age ($P < .001$); sex ($P = .005$); relative location of the 2 primary tumors ($P < .001$); diagnostic interval ($P < .001$); and IPLC characteristics such as year of diagnosis ($P = .004$), histology ($P < .001$), surgical procedure ($P < .001$), radiotherapy ($P = .04$), and chemotherapy for MSPLC ($P < .001$; Table 2). Figure 1A shows the survival outcomes among the 4 treatment groups ($P < .001$). Patients who only received radiotherapy had worse survival than those who underwent surgical resection but better survival than the no treatment group.

To evaluate the role of radiotherapy in terms of treatment for MSPLC, multiple PSM analyses were performed to compare radiotherapy with no treatment, surgery, and surgery plus radiotherapy. After PSM (ratio: 1:1; caliper=0.01), all baseline characteristics were matched well between the corresponding comparison groups (Table 2 and Tables S1 and S2 in Multimedia Appendix 1). As shown in Figure 1, the radiotherapy group had significantly better survival outcomes than the no treatment group ($P < .001$; Figure 1B) but significantly worse survival outcomes than the surgery group ($P < .001$; Figure 1C). However, radiotherapy seemed to not improve the survival outcome among patients who received surgery for MSPLC ($P = .26$; Figure 1D).
Table 2. Comparison of baseline characteristics between surgery and radiotherapy for second primary lung cancer before and after propensity score matching (PSM).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Before PSM</th>
<th>After PSM</th>
<th>p value</th>
<th>Characteristic</th>
<th>Before PSM</th>
<th>After PSM</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Radiation (n=864)</td>
<td>Surgery (n=759)</td>
<td>&lt;.001</td>
<td></td>
<td>Radiation (n=470)</td>
<td>Surgery (n=470)</td>
<td>.55</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>63.9 (8.9)</td>
<td>62.1 (9.0)</td>
<td>&lt;.001</td>
<td></td>
<td>63.0 (8.8)</td>
<td>62.7 (9.1)</td>
<td>.55</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
<td>.75</td>
</tr>
<tr>
<td>White</td>
<td>737 (85.3)</td>
<td>642 (84.6)</td>
<td></td>
<td>Black</td>
<td>401 (85.3)</td>
<td>393 (83.6)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>85 (9.8)</td>
<td>63 (8.3)</td>
<td></td>
<td>Other</td>
<td>39 (8.3)</td>
<td>45 (9.6)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>42 (4.9)</td>
<td>54 (7.1)</td>
<td></td>
<td></td>
<td>30 (6.4)</td>
<td>32 (6.8)</td>
<td></td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td>.005</td>
<td></td>
<td></td>
<td></td>
<td>.95</td>
</tr>
<tr>
<td>Male</td>
<td>417 (48.3)</td>
<td>313 (41.2)</td>
<td></td>
<td>Female</td>
<td>201 (42.8)</td>
<td>203 (43.2)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>447 (51.7)</td>
<td>446 (58.8)</td>
<td></td>
<td></td>
<td>269 (57.2)</td>
<td>267 (56.8)</td>
<td></td>
</tr>
<tr>
<td>Relative location, n (%)</td>
<td></td>
<td></td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
<td>.73</td>
</tr>
<tr>
<td>Ipsilateral</td>
<td>321 (37.2)</td>
<td>208 (27.4)</td>
<td></td>
<td>Contralateral</td>
<td>152 (32.3)</td>
<td>146 (31.1)</td>
<td></td>
</tr>
<tr>
<td>Contralateral</td>
<td>543 (62.8)</td>
<td>551 (72.6)</td>
<td></td>
<td></td>
<td>318 (67.7)</td>
<td>324 (68.9)</td>
<td></td>
</tr>
<tr>
<td>Diagnostic interval (months), mean (SD)</td>
<td>104.4 (48.7)</td>
<td>95.8 (45.3)</td>
<td>&lt;.001</td>
<td>99.1 (43.5)</td>
<td>100.9 (50.5)</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>IPLC&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year of diagnosis</td>
<td></td>
<td></td>
<td>.004</td>
<td></td>
<td></td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>1988-1995</td>
<td>242 (28)</td>
<td>256 (33.7)</td>
<td></td>
<td>1996-2003</td>
<td>313 (36.2)</td>
<td>286 (37.7)</td>
<td></td>
</tr>
<tr>
<td>2004-2012</td>
<td>309 (35.8)</td>
<td>217 (28.6)</td>
<td></td>
<td></td>
<td>161 (34.3)</td>
<td>166 (35.3)</td>
<td></td>
</tr>
<tr>
<td>SEER&lt;sup&gt;b&lt;/sup&gt; stage</td>
<td></td>
<td></td>
<td>.499</td>
<td></td>
<td></td>
<td></td>
<td>.73</td>
</tr>
<tr>
<td>Localized</td>
<td>560 (64.8)</td>
<td>505 (66.5)</td>
<td></td>
<td>1988-1995</td>
<td>242 (28)</td>
<td>256 (33.7)</td>
<td></td>
</tr>
<tr>
<td>Regional</td>
<td>304 (35.2)</td>
<td>254 (33.5)</td>
<td></td>
<td>1996-2003</td>
<td>313 (36.2)</td>
<td>286 (37.7)</td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
<td>.82</td>
</tr>
<tr>
<td>ADC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>451 (52.2)</td>
<td>495 (65.2)</td>
<td></td>
<td>1988-1995</td>
<td>242 (28)</td>
<td>256 (33.7)</td>
<td></td>
</tr>
<tr>
<td>SCC&lt;sup&gt;d&lt;/sup&gt;</td>
<td>280 (32.4)</td>
<td>173 (22.8)</td>
<td></td>
<td>1996-2003</td>
<td>313 (36.2)</td>
<td>286 (37.7)</td>
<td></td>
</tr>
<tr>
<td>Other NSCLC&lt;sup&gt;e&lt;/sup&gt;</td>
<td>133 (15.4)</td>
<td>91 (12)</td>
<td></td>
<td>2004-2012</td>
<td>309 (35.8)</td>
<td>217 (28.6)</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td>≥.99</td>
</tr>
<tr>
<td>Well differentiated</td>
<td>82 (9.5)</td>
<td>106 (14)</td>
<td></td>
<td>2004-2012</td>
<td>309 (35.8)</td>
<td>217 (28.6)</td>
<td></td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>295 (34.1)</td>
<td>259 (34.1)</td>
<td></td>
<td></td>
<td>161 (34.3)</td>
<td>166 (35.3)</td>
<td></td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>295 (34.1)</td>
<td>231 (30.4)</td>
<td></td>
<td></td>
<td>149 (31.7)</td>
<td>148 (31.5)</td>
<td></td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>42 (4.9)</td>
<td>33 (4.3)</td>
<td></td>
<td></td>
<td>22 (4.7)</td>
<td>23 (4.9)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>150 (17.4)</td>
<td>130 (17.1)</td>
<td></td>
<td></td>
<td>82 (14.9)</td>
<td>79 (16.8)</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
<td>.98</td>
</tr>
<tr>
<td>No surgery</td>
<td>102 (11.8)</td>
<td>47 (6.2)</td>
<td></td>
<td>2004-2012</td>
<td>309 (35.8)</td>
<td>217 (28.6)</td>
<td></td>
</tr>
<tr>
<td>Sublevel resection</td>
<td>100 (11.6)</td>
<td>105 (13.8)</td>
<td></td>
<td></td>
<td>149 (31.7)</td>
<td>148 (31.5)</td>
<td></td>
</tr>
<tr>
<td>Lobectomy</td>
<td>616 (71.3)</td>
<td>591 (77.9)</td>
<td></td>
<td></td>
<td>353 (75.1)</td>
<td>355 (75.5)</td>
<td></td>
</tr>
<tr>
<td>Pneumonectomy</td>
<td>46 (5.3)</td>
<td>16 (2.1)</td>
<td></td>
<td></td>
<td>13 (2.8)</td>
<td>12 (2.6)</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
<td>.87</td>
</tr>
<tr>
<td>Yes</td>
<td>155 (17.9)</td>
<td>131 (17.3)</td>
<td></td>
<td></td>
<td>91 (19.4)</td>
<td>88 (18.7)</td>
<td></td>
</tr>
<tr>
<td>No/unknown</td>
<td>709 (82.1)</td>
<td>628 (82.7)</td>
<td></td>
<td></td>
<td>379 (80.6)</td>
<td>382 (81.3)</td>
<td></td>
</tr>
<tr>
<td>Radiotherapy</td>
<td></td>
<td></td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>Characteristic</td>
<td>Before PSM</td>
<td>After PSM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Radiation (n=864)</td>
<td>Surgery (n=759)</td>
<td>P value</td>
<td>Radiation (n=470)</td>
<td>Surgery (n=470)</td>
<td>P value</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>176 (20.4)</td>
<td>123 (16.2)</td>
<td></td>
<td>86 (18.3)</td>
<td>88 (18.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No/unknown</td>
<td>688 (79.6)</td>
<td>636 (83.8)</td>
<td></td>
<td>384 (81.7)</td>
<td>382 (81.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SPLC**

<table>
<thead>
<tr>
<th>Chemotherapy</th>
<th>&lt;.001</th>
<th>≥.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>318 (36.8)</td>
<td>91 (12)</td>
</tr>
<tr>
<td>No/unknown</td>
<td>546 (63.2)</td>
<td>668 (88)</td>
</tr>
</tbody>
</table>

---

**Radiotherapy Versus Wedge Resection or Lobectomy**

To further compare survival between radiotherapy and specific surgical procedures, patients with MSPLC diagnosed with IPLC after 2004 were selected. Those who underwent unknown or indefinite sublevel resection, segmentectomy (very few patients) and pneumonectomy for MSPLC were excluded. There were 716 patients included for further analyses. The demographic characteristics are described in Table 3. Before PSM, Figure 2A shows that patients who underwent wedge resection or lobectomy had significantly better OS than those who received radiotherapy, and all of them had significantly better OS than the no treatment group. More clinical parameters such as T and N stage for IPLC and tumor size for MSPLC were matched by PSM, and all parameters were matched well (Tables S3-S5 in Multimedia Appendix 1). Similarly, after PSM, both wedge resection (P=.004; Figure 2C) and lobectomy (P=.002; Figure 2D) had significantly better OS than radiotherapy. Furthermore, radiotherapy also had greater survival benefits than no treatment (P<.001; Figure 2B).

---

*a*IPLC: initial primary lung cancer.

*b*SEER: Surveillance, Epidemiology, and End Results.

*c*ADC: adenocarcinoma.

*d*SCC: squamous cell carcinoma.

*e*NSCLC: non-small cell lung cancer.

*f*SPLC: second primary lung cancer.

Figure 1. (a) Overall survival of 2451 patients with MSPLC between 1988 and 2012 in different treatment groups before propensity score matching (PSM). (b) Overall survival of radio-therapy and none-treatment after PSM. (c) Overall survival of radiotherapy and surgery after PSM. (d) Overall survival of surgery and surgery plus radiotherapy after PSM.
Table 3. Demographic and clinical characteristics of 716 patients diagnosed with second primary lung cancer after 2004.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>65.8 (9.0)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>608 (84.9)</td>
</tr>
<tr>
<td>Black</td>
<td>65 (9.1)</td>
</tr>
<tr>
<td>Other</td>
<td>43 (6)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>310 (43.3)</td>
</tr>
<tr>
<td>Female</td>
<td>406 (56.7)</td>
</tr>
<tr>
<td><strong>Relative location, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Ipsilateral</td>
<td>279 (39)</td>
</tr>
<tr>
<td>Contralateral</td>
<td>437 (61)</td>
</tr>
<tr>
<td><strong>Interval, mean (SD)</strong></td>
<td>74.2 (21.4)</td>
</tr>
<tr>
<td><strong>Initial primary lung cancer</strong></td>
<td></td>
</tr>
<tr>
<td><strong>T stage, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>315 (44)</td>
</tr>
<tr>
<td>T2</td>
<td>277 (38.7)</td>
</tr>
<tr>
<td>T3</td>
<td>35 (4.9)</td>
</tr>
<tr>
<td>T4</td>
<td>66 (9.2)</td>
</tr>
<tr>
<td>Unknown</td>
<td>23 (3.2)</td>
</tr>
<tr>
<td><strong>N stage, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>528 (73.7)</td>
</tr>
<tr>
<td>N1</td>
<td>80 (11.2)</td>
</tr>
<tr>
<td>N2</td>
<td>97 (13.5)</td>
</tr>
<tr>
<td>Unknown</td>
<td>11 (1.5)</td>
</tr>
<tr>
<td><strong>Histology, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>ADC&lt;sup&gt;a&lt;/sup&gt;</td>
<td>406 (56.7)</td>
</tr>
<tr>
<td>SCC&lt;sup&gt;b&lt;/sup&gt;</td>
<td>206 (28.8)</td>
</tr>
<tr>
<td>Other NSCLC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>104 (14.5)</td>
</tr>
<tr>
<td><strong>Grade, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>94 (13.1)</td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>275 (38.4)</td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>211 (29.5)</td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>18 (2.5)</td>
</tr>
<tr>
<td>Unknown</td>
<td>118 (16.5)</td>
</tr>
<tr>
<td><strong>Surgery, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No surgery</td>
<td>131 (18.3)</td>
</tr>
<tr>
<td>Sublevel resection</td>
<td>106 (14.8)</td>
</tr>
<tr>
<td>Lobectomy</td>
<td>455 (63.5)</td>
</tr>
<tr>
<td>Pneumonection</td>
<td>24 (3.4)</td>
</tr>
<tr>
<td><strong>Chemotherapy, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>237 (33.1)</td>
</tr>
</tbody>
</table>

<sup>a</sup> ADC: adenocarcinoma; <sup>b</sup> SCC: squamous cell carcinoma; <sup>c</sup> NSCLC: non-small cell lung cancer.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>No/unknown</td>
<td>479 (66.9)</td>
</tr>
<tr>
<td><strong>Radiotherapy, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>174 (24.3)</td>
</tr>
<tr>
<td>No/unknown</td>
<td>542 (75.7)</td>
</tr>
<tr>
<td><strong>Second primary lung cancer</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Size (cm), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>0-3</td>
<td>385 (53.8)</td>
</tr>
<tr>
<td>3-5</td>
<td>71 (9.9)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>56 (7.8)</td>
</tr>
<tr>
<td>Unknown</td>
<td>204 (28.5)</td>
</tr>
<tr>
<td><strong>Surgery, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No surgery</td>
<td>533 (74.4)</td>
</tr>
<tr>
<td>Wedge resection</td>
<td>102 (14.2)</td>
</tr>
<tr>
<td>Lobectomy</td>
<td>81 (11.3)</td>
</tr>
<tr>
<td><strong>Chemotherapy, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>201 (28.1)</td>
</tr>
<tr>
<td>No/unknown</td>
<td>515 (71.9)</td>
</tr>
<tr>
<td><strong>Radiotherapy, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>309 (43.2)</td>
</tr>
<tr>
<td>No/unknown</td>
<td>407 (56.8)</td>
</tr>
<tr>
<td><strong>Treatment, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>224 (31.3)</td>
</tr>
<tr>
<td>Only radiation</td>
<td>309 (43.2)</td>
</tr>
<tr>
<td>Only wedge</td>
<td>102 (14.2)</td>
</tr>
<tr>
<td>Only lobectomy</td>
<td>81 (11.3)</td>
</tr>
</tbody>
</table>

aADC: adenocarcinoma.
bSCC: squamous cell carcinoma.
cNSCLC: non-small cell lung cancer.
Figure 2. Overall survival of (A) 716 patients with metachronous second primary lung cancer (MSPLC) after 2004 in different treatment groups before propensity score matching (PSM); (B) patients who received radiotherapy or no treatment, after PSM; (C) patients who received radiotherapy or underwent wedge resection, after PSM; (D) patients who received radiotherapy or underwent lobectomy, after PSM.

ML-Based Cancer-Specific Death Risk Prediction

Using LASSO regression, we identified 9 variables that made significant contributions to CSS (Figure 3). These variables encompassed age at diagnosis, sex, year of diagnosis, radiotherapy of IPLC, primary site, histology, surgery, chemotherapy, and radiotherapy of MPSLC. The ML models displayed outstanding performance, as indicated by high AUC values, highlighting the superiority of artificial intelligence in prognostic prediction (Figure 4). The decision curve analyses are depicted in Figure 5. Additionally, we assessed the sensitivity and specificity of each ML model using the maximal Youden index, which represents an optimal balance between true positives and true negatives (Table 4). Through 5-fold cross-validation, the XGB, RFC, and ADB models demonstrated superior performance. In order to gain deeper insights into the relationships between demographic characteristics and long-term outcomes for MSPLC patients, we used these ML algorithms to develop predictive models to assess the 1-year, 3-year, 5-year, and 10-year risks of cumulative cancer-specific mortality based on the aforementioned variables. Consequently, we calculated the contribution of each variable. Notably, we identified the variables associated with CSS at different time intervals (Figure 6). Surgery for MPSLC predominantly and substantially influenced 1-year, 3-year, 5-year, and 10-year survival, but its effect was comparatively less than that of surgery. The primary site and histology of MPSLC affected 1-year, 3-year, and 5-year CSS, but it had no impact on 10-year CSS. Radiotherapy for IPLC also had an impact on 1-year and 3-year CSS but had minimal influence on 5-year and 10-year survival.
Figure 3. Machine learning model using least absolute shrinkage and selection operator (LASSO) regression analysis for risk prediction of cumulative cancer-specific mortality in patients with metachronous second primary lung cancer (MSPLC): (A) 5-fold cross-validation results and (B) model regression coefficient profile.
Figure 4. Receiver operating characteristic (ROC) curves for machine learning models for risk prediction of cumulative cancer-specific mortality in patients with metachronous second primary lung cancer (MSPLC): (A) 1-year lymphoma-specific mortality; (B) 3-year lymphoma-specific mortality; (C) 5-year lymphoma-specific mortality; (D) 10-year lymphoma-specific mortality. ADB: adaptive boosting; ANN: artificial neural network; AUC: area under the curve; GBDT: gradient boosting decision tree; KNN: K nearest neighbor; RFC: random forest classifier; ROC: receiver operating characteristic; XGB: extreme gradient boosting.

Figure 5. Decision curve analysis for 6 classical machine learning–based models for risk prediction of cumulative cancer-specific mortality in patients with metachronous second primary lung cancer (MSPLC): (A) 1-year lymphoma-specific mortality; (B) 3-year lymphoma-specific mortality; (C) 5-year lymphoma-specific mortality; (D) 10-year lymphoma-specific mortality.

<table>
<thead>
<tr>
<th>Model</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>AUC(^a) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-year cancer-specific survival</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XGB(^b)</td>
<td>77</td>
<td>60.2</td>
<td>0.73 (0.71-0.75)</td>
</tr>
<tr>
<td>RFC(^c)</td>
<td>76.7</td>
<td>63</td>
<td>0.74 (0.72-0.76)</td>
</tr>
<tr>
<td>ADB(^d)</td>
<td>83.1</td>
<td>54.4</td>
<td>0.75 (0.73-0.77)</td>
</tr>
<tr>
<td>KNN(^e)</td>
<td>70.9</td>
<td>63.6</td>
<td>0.72 (0.70-0.74)</td>
</tr>
<tr>
<td>ANN(^f)</td>
<td>88.2</td>
<td>41.9</td>
<td>0.74 (0.72-0.76)</td>
</tr>
<tr>
<td>GBDT(^g)</td>
<td>90.6</td>
<td>36</td>
<td>0.74 (0.72-0.76)</td>
</tr>
<tr>
<td><strong>3-year cancer-specific survival</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XGB</td>
<td>69.9</td>
<td>73.8</td>
<td>0.77 (0.75-0.79)</td>
</tr>
<tr>
<td>RFC</td>
<td>75.6</td>
<td>69.2</td>
<td>0.77 (0.75-0.79)</td>
</tr>
<tr>
<td>ADB</td>
<td>79.3</td>
<td>66.4</td>
<td>0.76 (0.74-0.78)</td>
</tr>
<tr>
<td>KNN</td>
<td>79.6</td>
<td>64</td>
<td>0.75 (0.73-0.77)</td>
</tr>
<tr>
<td>ANN</td>
<td>83.6</td>
<td>59.9</td>
<td>0.77 (0.75-0.79)</td>
</tr>
<tr>
<td>GBDT</td>
<td>84.4</td>
<td>57.6</td>
<td>0.75 (0.73-0.77)</td>
</tr>
<tr>
<td><strong>5-year cancer-specific survival</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XGB</td>
<td>79.6</td>
<td>71.3</td>
<td>0.78 (0.75-0.81)</td>
</tr>
<tr>
<td>RFC</td>
<td>79.2</td>
<td>71.5</td>
<td>0.79 (0.76-0.82)</td>
</tr>
<tr>
<td>ADB</td>
<td>75.3</td>
<td>74.7</td>
<td>0.79 (0.76-0.82)</td>
</tr>
<tr>
<td>KNN</td>
<td>74.3</td>
<td>73.9</td>
<td>0.77 (0.74-0.80)</td>
</tr>
<tr>
<td>ANN</td>
<td>79.3</td>
<td>71.5</td>
<td>0.80 (0.77-0.83)</td>
</tr>
<tr>
<td>GBDT</td>
<td>80.1</td>
<td>69.5</td>
<td>0.78 (0.75-0.81)</td>
</tr>
<tr>
<td><strong>10-year cancer-specific survival</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XGB</td>
<td>78.8</td>
<td>74.7</td>
<td>0.84 (0.80-0.88)</td>
</tr>
<tr>
<td>RFC</td>
<td>78.3</td>
<td>40.7</td>
<td>0.83 (0.79-0.87)</td>
</tr>
<tr>
<td>ADB</td>
<td>78.4</td>
<td>81</td>
<td>0.84 (0.80-0.88)</td>
</tr>
<tr>
<td>KNN</td>
<td>80.7</td>
<td>73.4</td>
<td>0.78 (0.72-0.84)</td>
</tr>
<tr>
<td>ANN</td>
<td>68.8</td>
<td>88.6</td>
<td>0.85 (0.81-0.89)</td>
</tr>
<tr>
<td>GBDT</td>
<td>79.7</td>
<td>78.5</td>
<td>0.85 (0.81-0.89)</td>
</tr>
</tbody>
</table>

\(^a\)AUC: area under the curve.  
\(^b\)XGB: extreme gradient boosting.  
\(^c\)RFC: random forest classifier.  
\(^d\)ADB: adaptive boosting.  
\(^e\)KNN: K nearest neighbor.  
\(^f\)ANN: artificial neural network.  
\(^g\)GBDT: gradient boosting decision tree.
Age-Adjusted Competing Risk Analysis

To gain further insights into the cumulative incidence associated with each variable, we conducted competing risk analyses (Figure 7). Female MPSLC patients had lower cumulative cancer-specific mortality (hazard ratio [HR]=0.79, 95% CI 0.71-0.87; \(P<.001\)). Patients diagnosed with IPLC in more recent years also had lower cumulative cancer-specific mortality: 1996-2003 (HR=0.85, 95% CI 0.76-0.96; \(P<.001\)); 2004-2012 (HR=0.79, 95% CI 0.73-0.85; \(P<.001\)). However, patients who received radiotherapy for their IPLC had increased mortality (HR=1.31, 95% CI 1.16-1.50; \(P<.001\)). The histology of the second primary lung cancer played a significant role, with higher mortality rates for squamous carcinoma than adenocarcinoma (HR=1.28, 95% CI 1.12-1.46; \(P<.001\)). Moreover, the use of surgery for the second primary lung cancer was associated with lower mortality rates. This was particularly true for sublevel resection (HR=0.37, 95% CI 0.32-0.43; \(P<.001\)), lobectomy (HR=0.56, 95% CI 0.51-0.61; \(P<.001\)), and pneumonectomy (HR=0.80, 95% CI 0.71-0.89; \(P<.001\)). Conversely, the use of chemotherapy or radiotherapy for the second primary lung cancer was associated with increased mortality rates, potentially due to the severity of the patients’ initial condition (chemotherapy: HR=1.64, 95% CI 1.47-1.83; \(P<.001\); radiotherapy: HR=1.19, 95% CI 1.07-1.33; \(P<.001\)). Therefore, we performed additional analyses for the different treatment modalities, as shown in Figure 8. Surgery alone (HR=0.83, 95% CI 0.81-0.85; \(P<.001\)) had the lowest cancer-specific mortality, followed by surgery and chemotherapy (HR=0.76, 95% CI 0.71-0.82; \(P<.001\)) and surgery and radiotherapy (HR=0.79, 95% CI 0.71-0.88; \(P<.001\)). Among different surgical approaches, sublevel resection alone (HR=0.26, 95% CI 0.21-0.31, \(P<.001\)) had the lowest mortality rate, followed by pneumonectomy alone (HR=0.72, 95% CI 0.63-0.82, \(P<.001\)) and lobectomy alone (HR=0.92, 95% CI 0.78-1.09; \(P<.001\)).
Figure 7. Cumulative cancer-specific mortality per age-adjusted competing risk analysis in subgroup analysis by (A) age, (B) sex, (C) year of diagnosis of the initial primary lung cancer, (D) radiotherapy of the initial primary lung cancer, (E) primary site of the second primary lung cancer, (F) histology of the second primary lung cancer, (G) surgery for the second primary lung cancer, (H) chemotherapy of the second primary lung cancer, (I) radiotherapy of the second primary lung cancer. HR: hazard ratio.
Figure 8. Age-adjusted competing risk analysis to estimate the cumulative cancer-specific mortality with different treatment modalities: (A) treatment for second primary lung cancer, (B) radiotherapy alone versus surgery alone. HR: hazard ratio.

Discussion

Principal Findings

With the rapid advancement and wide application of low-dose computed tomography for screening of pulmonary nodules, more patients are being diagnosed with MSPLC. Multiple primary lung cancer (MPLC) is a special kind of lung carcinoma that can be categorized into synchronous MPLC and metachronous MPLC. Among these, MSPLC is the most common form of MPLC that can be expected to receive curable management. However, limited progress has been made so far on accurate diagnoses, optimal medical interventions, and prognostic outcomes. In this study, our findings suggest that surgical resections, including wedge resection and lobectomy, contribute to better survival rates than radiation therapy in the context of MSPLC. However, it is important to note that radiation therapy remains a viable and valid alternative for the treatment of MSPLC.

Surgical resection is reportedly feasible for MSPLC and could significantly improve the prognosis [5-9], but the role of radiation therapy in the treatment of MSPLC remains unclear. Considering that patients who previously underwent surgery for IPLC may not tolerate another pulmonary resection, finding optimal alternative treatments is important. Therefore, using the population-based SEER database, this study used PSM analyses and ML techniques to first compare survival outcomes between patients who received radiotherapy or underwent surgical resection for MSPLC.

Of all enrolled patients, most (2187/2451, 89.2%) had undergone surgery for IPLC before (Table 1). However, 65.4% (1603/2451) of the patients with MSPLC did not undergo surgical resection for MSPLC, and 35.3% (864/2451) of them received radiation therapy (Table 1). It could be inferred that a considerable proportion of patients with MSPLC could not tolerate another surgical resection, and radiotherapy might be the predominant alternative treatment for them. Although surgical resection was first recommended for patients with MSPLC, radiation therapy is also important, especially for inoperable cancers. Given that very few studies have focused on long-term survival outcomes after radiotherapy versus surgery, this study may provide a more solid indication in terms of the use of radiotherapy for patients with MSPLC.

Previous studies reported that 5-year OS rates for patients with MSPLC varied, ranging from 26% to 38% [20-22]; these rates are similar to that of our study (34.7% for the entire cohort). The 5-year survival rates were 18.0% for radiotherapy, 49.3% for surgery, 38.8% for surgery plus radiotherapy, and 7.7% for no treatment. Ono et al [13] reported on 19 patients who were diagnosed with MSPLC after lung resection for IPLC and underwent proton beam therapy. Their research showed a 3-year survival rate of 63.2% and a 3-year local control rate of 84.2%, which indicated the safety and feasibility of proton beam therapy for patients with MSPLC. Miyazaki et al [23] compared survival
outcomes among metachronous MPLC patients after stereotactic body radiotherapy (N=26) and surgery (N=51) and found no significant differences. The study by Taioli et al [24] included 494 cases from the SEER database and showed that OS was better with surgery than with radiation therapy after the treatment of MSPLC [24]. However, their inclusion criteria were not rigorous enough; the diagnostic interval between the 2 primary lung tumors was too short (6 months), which could fail to exclude patients with relapse or metastasis. Additionally, the analyses were not adjusted for confounding factors, and this might have caused significant bias. In our study, multiple PSM analyses were performed to control for confounding effects. The surgery group had significantly better survival than the radiotherapy group (P<.001), and the radiotherapy group had greater survival than the no treatment group (P<.001; Figure 1). Therefore, surgical resection should be considered first for patients with MSPLC if their physical condition and pulmonary function reserve permit. For those with an inoperable cancer or who are not willing to undergo another surgery, radiation therapy may be an alternative. Additionally, after PSM, there was no significant difference between the surgery and surgery plus radiotherapy groups (P=.26), which indicated that preoperative or postoperative radiotherapy might not increase survival benefits for patients with MSPLC.

Lobectomy remains the commonly accepted standard treatment for resectable NSCLC. In recent years, sublobar resections have been widely reported to be adequate in early-stage NSCLC, resulting in less impairment and greater respiratory function reserve [1,25,26]. However, the prognostic role of sublobar resection among patients with MSPLC has not been clearly clarified. Yang et al [8] identified 454 matched pairs of patients with MSPLC receiving lobectomy or sublobar resection from the SEER database and found that the lobectomy group had significantly better survival than the sublobar resection group. Lee et al [27] concluded that MSPLC had similar survival outcomes with wedge resection and lobectomy by analyzing 625 patients with a diagnostic interval ≥6 months. There have been few studies that have focused on survival outcomes after radiotherapy compared with wedge resection or lobectomy. Thus, to further verify the rigor of our study, patients diagnosed after 2004 and with definite therapeutic information (only including no treatment, radiotherapy, wedge resection, and lobectomy) were selected to compare survival outcomes between those undergoing radiotherapy or specific surgical resections. Very few cases underwent segmentectomy (n=16) or surgery plus radiotherapy (n=11) for MSPLC and were excluded. Additionally, T and N stage (American Joint Committee on Cancer, 6th edition) for IAPLC and tumor size were also adjusted using PSM analyses. Of the 716 patients diagnosed after 2004, 53.8% had MSPLC with a tumor size ≤3 cm, while a limited number of patients (127/716, 17.3%; Table 2) had a tumor larger than 3 cm, though some patients’ MSPLC tumor sizes were unknown (204/716, 28.5%). This implied that most of the patients with MSPLC could be categorized as “early-stage” NSCLC if their tumors were recorded as initial lung cancer, which is a strong indication for sublobar resection and radiation therapy. There were actually only a few patients that underwent lobectomy for MSPLC (entire sample: 352/2451, 14.4%; diagnosed after 2004: 81/716, 11.3%). Radiotherapy seemed to be the most common treatment for MSPLC (entire sample: 864/2451, 35.3%; diagnosed after 2004: 309/716, 43.2%). All the aforementioned facts indicate that most patients with MSPLC might not tolerate another surgical resection, especially lobectomy, or be more willing to receive noninvasive radiation therapy. Therefore, comparing survival outcomes between radiotherapy and wedge resection or lobectomy is highly necessary. As shown in Figure 2, the radiotherapy group also had significantly greater OS than the no treatment group (P<.001) but poorer OS than both the lobectomy (P=.002) and wedge resection (P=.004) groups. When patients’ physical condition and pulmonary function reserve permit, whether choosing lobectomy or wedge resection, patients undergoing surgical resection may gain greater survival benefits than those receiving radiation therapy.

The development of long-term outcome prediction models using ML techniques represents a significant breakthrough in the field of MSPLC. This paper convincingly demonstrates the utility of ML algorithms for accurately predicting cumulative cancer-specific mortality at various time intervals. The exceptional performance of these predictive models emphasizes the superiority of artificial intelligence in prognostic prediction, offering precise and reliable predictions for individual patients. Integrating such models into routine clinical practice has the potential to optimize treatment strategies and improve patient outcomes in MSPLC. Furthermore, the study uses competing risk analysis to delve into the impact of different factors on CSS among MSPLC patients across distinct time intervals. The findings provide valuable insights into the factors influencing both short-term (1-year and 3-year) and long-term (5-year and 10-year) survival outcomes. This enhanced understanding of the factors affecting patient outcomes contributes to improved prognostic assessments and facilitates informed treatment decision-making by clinicians.

Generally, patients with MLC had better survival outcomes than those with intrapulmonary metastases from IPLC after surgery [22,28]. However, effective methods to accurately identify MPLC patients have not existed until now. Previous studies identified patients with MSPLC using inclusion and exclusion criteria that lacked rigor [8,24,27]. In this study, to avoid the potential confounding effect of metastases, we only included patients with a diagnostic interval greater than 4 years, which indicated a thoroughly representative group of patients with MSPLC [16].

To the best of our knowledge, using PSM analyses and ML techniques on the largest cohort of patients with MSPLC, this study is the first to compare the survival outcomes after radiotherapy with those after surgical resection for MSPLC. Nevertheless, limitations in some aspects of the study still exist. First, this is a retrospective study based on the study population from the SEER database. A certain degree of data bias could not be totally avoided. Second, there might have been an inclination for treatment regarding the patients who received radiotherapy, because they were usually ineligible for surgery due to poorer physical condition and insufficient pulmonary function reserve. Thus, though we tried to control for the confounding effects using PSM, patient bias between different treatment groups also existed because details on physical
condition and lung function were unknown. Further evaluation should be performed by prospective studies in the future. Third, since very few patients underwent a pneumonectomy and were thus excluded from our study, the prognostic role of pneumonectomy for patients with MSPLC requires a large cohort to verify. Additionally, we acknowledge the limitations inherent in the SEER database, which lacks comprehensive information, including details on immunotherapy and targeted therapy and the specifics of radiotherapy such as the target volume, treatment dose, and radiation technology. We hope that future cohort studies will incorporate these specifics to provide a more comprehensive understanding of the treatment landscape for MSPLC.

Conclusions

Overall, this study indicated that surgical resections such as wedge resection and lobectomy performed better than radiation therapy in terms of survival of patients with MSPLC. However, many patients with MSPLC may not tolerate surgery because of previously treated initial lung cancer. Among the treatment options, radiation therapy confers great survival outcomes and can be a valid alternative for surgery. Future prospective studies can be designed to further confirm the effectiveness of radiation therapy for MSPLC.

Acknowledgments

This work was supported by the Postdoctoral Fellowship Program of CPSF (GZB20230481), Post-Doctor Research Project, West China Hospital, Sichuan University (2024HXBH149, 2024HXBH006), National Natural Science Foundation of China (82303773, 82303772, 82303694), Natural Science Foundation of Sichuan Province (2023NSFSC1885, 024NSFSC1908), and Key Research and Development Program of Sichuan Province (23ZDYF2836).

Data Availability

The data sets generated and analyzed during the current study are available in the SEER database [14].

Authors' Contributions

YW conceptualized and supervised the study, acquired the funding, validated the data, and reviewed and edited the manuscript draft. YZ curated the data. YZ, AZ, and YY performed the methodology and wrote the original manuscript draft. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplemental tables.

References


Abbreviations

ADB: adaptive boosting
ANN: artificial neural network
AUC of ROCs: area under the receiver operating characteristic curves
CSS: cancer-specific survival
GBDT: gradient boosting decision tree
HR: hazard ratio
ICD-O: International Classification of Diseases for Oncology
IPLC: initial primary lung cancer
KNN: K nearest neighbor
LASSO: least absolute shrinkage and selection operator
ML: machine learning
MPLC: multiple primary lung cancer
MSPLC: metachronous second primary lung cancer
NSCLC: non-small cell lung cancer
OS: overall survival
PSM: propensity score matching
RFC: random forest classifier
SEER: Surveillance, Epidemiology, and End Results
WHO: World Health Organization
XGB: extreme gradient boosting

Please cite as:
JMIR Cancer 2024;10:e53354
URL: https://cancer.jmir.org/2024/1/e53354
doi:10.2196/53354
PMID:38865182
Patterns of Prescription Medication Use Before Diagnosis of Early Age-Onset Colorectal Cancer: Population-Based Descriptive Study

Vienna Cheng¹,², PharmD; Eric C Sayre³, PhD; Vicki Cheng¹,², PharmD, MSc; Ria Garg⁴, PharmD; Sharlene Gill⁵,⁶, MPH, MBA, MD; Ameer Farooq⁷, MPH, MD; Mary A De Vera¹,²,⁸, BSc, MSc, PhD
¹Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada
²Collaboration for Outcomes Research and Evaluation, Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada
³British Columbia Centre on Substance Use, Vancouver, BC, Canada
⁴Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada
⁵Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
⁶BC Cancer, Vancouver, BC, Canada
⁷School of Medicine Division of General Surgery, Queen’s University, Kingston, ON, Canada
⁸Centre for Health Evaluation and Outcome Sciences, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Mary A De Vera, BSc, MSc, PhD
Faculty of Pharmaceutical Sciences
University of British Columbia
2405 Wesbrook Mall
Vancouver, BC, V6T 1Z3
Canada
Phone: 1 6048272138
Email: mdevera@mail.ubc.ca

Abstract

Background: Colorectal cancer (CRC) is estimated to be the fourth most common cancer diagnosis in Canada (except for nonmelanoma skin cancers) and the second and third leading cause of cancer-related death in male and female individuals, respectively.

Objective: The rising incidence of early age-onset colorectal cancer (EAO-CRC; diagnosis at less than 50 years) calls for a better understanding of patients’ pathway to diagnosis. Therefore, we evaluated patterns of prescription medication use before EAO-CRC diagnosis.

Methods: We used linked administrative health databases in British Columbia (BC), Canada, to identify individuals diagnosed with EAO-CRC between January 1, 2010, and December 31, 2016 (hereinafter referred to as “cases”), along with cancer-free controls (1:10), matched by age and sex. We identified all prescriptions dispensed from community pharmacies during the year prior to diagnosis and used the Anatomical Therapeutic Chemical Classification system Level 3 to group prescriptions according to the drug class. A parallel assessment was conducted for individuals diagnosed with average age-onset CRC (diagnosis at age 50 years and older).

Results: We included 1001 EAO-CRC cases (n=450, 45% female participants; mean 41.0, SD 6.1 years), and 12,989 prescriptions were filled in the year before diagnosis by 797 (79.7%) individuals. Top-filled drugs were antidepressants (first; n=1698, 13.1%). Drugs for peptic ulcer disease and gastroesophageal reflux disease (third; n=795, 6.1%) were more likely filled by EAO-CRC cases than controls (odds ratio [OR] 1.4, 95% CI 1.2-1.7) and with more frequent fills (OR 1.8, 95% CI 1.7-1.9). We noted similar patterns for topical agents for hemorrhoids and anal fissures, which were more likely filled by EAO-CRC cases than controls (OR 7.4, 95% CI 5.8-9.4) and with more frequent fills (OR 15.6, 95% CI 13.1-18.6).

Conclusions: We observed frequent prescription medication use in the year before diagnosis of EAO-CRC, including for drugs to treat commonly reported symptoms of EAO-CRC.

(JMIR Cancer 2024;10:e50402) doi:10.2196/50402
KEYWORDS
colorectal cancer; medications; medication patterns; cancer diagnosis; prediagnosis; prescriptions; patterns; early-onset; population-based; incidence; male individuals; female individuals; health databases; pharmacology; diagnostic; descriptive study; gastroenterology; cancers

Introduction

Colorectal cancer (CRC) is estimated to be the fourth most common cancer diagnosis in Canada (except for nonmelanoma skin cancers) and the second and third leading cause of cancer-related death in male and female individuals, respectively [1]. Given the marked onset of CRC among individuals aged 50 years, it was historically considered a disease for older adults. However, recent evidence particularly over the past decade has revealed a rise in the incidence of early age-onset CRC (EAO-CRC), defined as diagnosis among those younger than 50 years [2]. For example, a 2020 Canadian study [3] showed that between 2008 and 2017, the 30- to 39-year-age group accounted for the most significant increase with age-specific average annual percent changes of 4.33 (95% CI 2.79-5.91) for female individuals and 4.53 (95% CI 2.89-6.19) for male individuals.

The increasing incidence of EAO-CRC has called for research to better understand various aspects of the disease [2-4], including the path to diagnosis, particularly patterns of health care use. Using administrative health databases in British Columbia (BC), Canada, a 2022 case-control study found that in comparison to age- and sex-matched cancer-free controls, individuals diagnosed with EAO-CRC experienced a marked increase in outpatient physician visits during the year prior to diagnosis, with the reason for visit most commonly documented as nausea, vomiting, and abdominal pain [5]. Therefore, delineating patterns of prescription medication use before diagnosis of EAO-CRC may provide further insight, particularly as certain pharmacologic treatments may suggest potential diagnostic opportunities for EAO-CRC. In 2017, Pottegård and Hallas [6] used the Danish Cancer Registry to evaluate prescription drug use in the 24 months preceding a diagnosis of lung, breast, colon, and prostate cancers and found a stable pattern that markedly increased at 6 months before diagnosis. Among a prespecified list of drug classes that may likely be prescribed for early symptoms of one of the cancers studied (eg, drugs against overactive bladder may be associated with future prostate cancer diagnosis and drugs against constipation or diarrhea may be associated with future colon cancer diagnosis), such as opioids, oral antidiabetics, and statins, authors found that for those with colon cancer, the increased prescription rates before diagnosis were for proton pump inhibitors and antibiotics [6]. It is important to assess whether a similar pattern is present in another jurisdiction with a specific focus on CRC and considering age at diagnosis, particularly given the increasing incidence of EAO-CRC [2-4]. Thus, our primary aim was to assess patterns of prescription medication use among individuals with EAO-CRC during the year preceding diagnosis. To contextualize our findings, we also assessed patterns of prescription medication use among age- and sex-matched cancer-free controls and individuals diagnosed with average-age onset CRC (AAO-CRC; 50 years and older). We aim to better understand the pathway to diagnosis through evaluating patterns of prescription medication use in the year preceding EAO-CRC diagnosis.

Methods

Data Sources

As with prior population-based research on the epidemiology of EAO-CRC [7], we linked administrative health databases capturing longitudinal and deidentified individual-level health services data for the province of BC, Canada [8-14]. Population Data BC facilitated data access to the Medical Services Plan database on outpatient visits [13], the Discharge Abstract Database on inpatient visits [14], the Consolidation File for demographics [11], the Vital Statistics File for deaths [12], and the PharmaNet database on all prescriptions dispensed in community pharmacies regardless of payer [15]. These databases were linked to the BC Cancer Registry, which includes data on cancer diagnosis (eg, date and site) [9].

Study Design

A population-based descriptive observational study was conducted. First, we identified CRC cases as individuals diagnosed with CRC between January 1, 2010, and December 31, 2016, using International Classification of Diseases for Oncology, Third Edition (ICD-O-3) codes, specifically: C18.2-C18.9 (colon), C19.9 (rectosigmoid), and C20 and C21.8 (rectum). Our study period coincided with the beginning (in 2010) of population-based reporting of staging data, based on American Joint Committee on Cancer staging guidelines, with >85% capture in the BC Cancer Registry [16,17]. We assigned the index date as the date of definitive diagnosis from the BC Cancer Registry based on tissue diagnosis of CRC (endoscopist, surgeon, or oncologist). Next, we further classified cases as those with EAO-CRC (diagnosed at less than 50 years of age) and AAO-CRC (diagnosed at 50 years of age or later). We matched individuals with CRC to cancer-free controls (1:up to 10) on age and sex. Controls were also required to have a health care use (ie, outpatient visit, hospitalization, or prescription fill) within the same year their matched case was diagnosed. Controls were assigned an index date, which corresponded to their match date (Multimedia Appendix 1 illustrates data sources and study sample).

Prescription Medication Use

We assessed the use of prescription medications over the 1-year period preceding the index date using the PharmaNet database. We drew rationale for evaluating the 1-year period before diagnosis from the study by Pottegård and Hallas [6] showing marked prescription drug use 6 months before cancer diagnosis and from our own prior work with patterns of outpatient physician visits the year before cancer diagnosis [5]. By law, prescriptions dispensed from community pharmacies in BC must be entered in PharmaNet, a province-wide network [15].
Thus, we were able to assess all prescriptions, regardless of payer, and extracted relevant information including prescription date, drug identification number, and Anatomical Therapeutic Chemical (ATC) classification [18]. In particular, we used the third-level ATC code, allowing us to categorize drugs according to first level—main anatomical or pharmacological group (eg, A alimentary tract and metabolism); second level—pharmacological or therapeutic subgroup (eg, A10 drugs used in diabetes); third and fourth levels—chemical, pharmacological, or therapeutic subgroup (eg, A10B blood glucose–lowering drugs and A10BA biguanides; Multimedia Appendix 2).

Statistical Analysis

We used descriptive statistics (eg, mean and proportions) to characterize all individuals included in our study sample according to age, sex (female or male), socioeconomic status (determined using neighborhood income per person equivalent adjusted for household size), type of residence (rural vs urban, determined using Census Metropolitan Area or Census Agglomeration from geographical census data). For individuals with CRC, the cancer site using ICD-O-3 codes (eg, rectum, left colon, right colon, and transverse colon) and stage at diagnosis were also determined.

We assessed patterns of prescriptions among EAO-CRC cases overall and according to sex and stage at diagnosis, reporting counts and proportions using both prescriptions and persons as units of analyses. Using logistic regression, we evaluated determinants of our outcome of having ≥1 prescription filled in the year before diagnosis among EAO-CRC cases. Potential determinants included age, sex, neighborhood income quintile, residence, cancer diagnosis site, and stage. We used a backward stepwise approach and retained the model variables based on statistical or clinical significance. We then compared patterns of prescription medications among EAO-CRC cases and controls, reporting counts, proportions, and odds ratios (ORs) and corresponding 95% CI, where relevant. We also compared patterns of prescription medications among EAO-CRC and AAO-CRC cases, reporting counts, proportions, ORs and corresponding 95% CIs, where relevant. We completed all these analyses using SAS statistical software (version 9.4; SAS Institute).

Study Conduct

All inferences, opinions, and conclusions drawn in this paper are those of the authors and do not reflect the opinions or policies of the Data Stewards.

Ethical Considerations

This study was approved by the University of British Columbia’s Behavioural Research Ethics Board (H17-03530) and was performed in accordance with relevant guidelines and regulations. Consent to participate was waived by the University of British Columbia’s Behavioural Research Ethics Board, as this research involves secondary use of data. Individual-level health services data from the linked administrative health databases were deidentified or scrambled.

Results

Our study included 1001 cases with EAO-CRC (n=450, 45% female participants; mean age 41.0, SD 6.1 years) and 10,010 matched cancer-free controls (n=4500, 45% female participants; mean age 41.0, SD 6.1 years). As shown in Table 1, EAO-CRC cases were most frequently diagnosed with cancer in the rectum (n=418, 41.8%) and with stage III (n=351, 35%) and stage IV (n=270, 27%) disease. In our parallel analyses, we identified 12,331 cases with AAO-CRC (n=5536, 44.9% female participants, mean age 66.6, SD 9.2 years), who were most frequently diagnosed with cancer in the left colon (n=5210, 42.3%) and stage III (n=3644, 29.6%) or stage II (n=2996, 24.3%) disease.

There were 12,989 prescription events among 797 (79.7%) EAO-CRC cases and 174,806 prescription events among 7796 (77.9%) matched cancer-free controls. With respect to individuals, there is no significant difference in the proportions of EAO-CRC cases and controls filling prescriptions (OR 1.11, 95% CI 0.94-1.3). However, with respect to the number of prescriptions filled, among 797 EAO-CRC cases, there was a mean of 16.3 (SD 73.7) prescriptions (median 5.0) per case; whereas for 7796 controls, there was a mean of 22.4 (SD 99.3) prescriptions (median 6.0) per control. Multimedia Appendix 3 summarizes medication classes that represent ≥1% of prescriptions for EAO-CRC cases and n≥1748 prescriptions for controls of all prescriptions in the year before diagnosis for EAO-CRC cases and controls. Assessing specific medications including ranking and frequency revealed patterns of use. For example, antidepressants (ATC3 N06A) were the top medications filled by both EAO-CRC cases (n=1698, 13.1% of prescriptions) and controls (n=17262, 9.9% of prescriptions) with EAO-CRC having more frequent fills (OR 1.4, 95% CI 1.3-1.4) than cases. Gastrointestinal drugs (ATC3 N02A; for peptic ulcer disease and gastrosophageal reflux disease) were the third most filled prescriptions by EAO-CRC cases (n=795, 6.1% of prescriptions) and fifth most filled by controls (n=6126, 3.5% of prescriptions) with EAO-CRC cases having higher odds of filling (OR 1.4, 95% CI 1.2-1.7) and having more frequent fills (OR 1.8, 95% CI 1.7-1.9). Relatedly, agents for the treatment of hemorrhoids and anal fissures for topical use (ATC3 C05A) and drugs for constipation (ATC3 A06A) represent the ninth (n=275, 2.1% of prescriptions) and tenth (n=250, 1.9% of prescriptions) most filled prescriptions by EAO-CRC cases, respectively, but were not among ≥1% (n≥1748 prescriptions) of prescriptions for controls. EAO-CRC cases had higher odds of filling (OR 7.4, 95% CI 5.8-9.4) and had more frequent fills (OR 15.6, 95% CI 13.1-18.6) for topical agents for hemorrhoids and anal fissures. Among EAO-CRC cases, factors associated with filling ≥1 or more prescriptions in the year before diagnosis included having inflammatory bowel disease (adjusted odds ratio [aOR] 3.43; 95% CI 1.20-9.78) and depression (aOR 4.20, 95% CI 1.49-11.85). As well, number of outpatient visits was also a determinant with an aOR of 1.14 (95% CI 1.09-1.18).
Table 1. Characteristics of individuals with EAO-CRC\(^a\) (less than 50 years), AAO-CRC\(^b\) (50 years and older), and their respective controls.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>EAO-CRC Cases (n=1001)</th>
<th>EAO-CRC Controls (n=10,010)</th>
<th>AAO-CRC Cases (n=12,331)</th>
<th>AAO-CRC Controls (n=123,310)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>41 (6.1)</td>
<td>41 (6.1)</td>
<td>66.6 (9.2)</td>
<td>66.6 (9.2)</td>
</tr>
<tr>
<td>Female participants, n (%)</td>
<td>450 (45)</td>
<td>4500 (45)</td>
<td>5536 (44.9)</td>
<td>55,360 (44.9)</td>
</tr>
<tr>
<td>Neighborhood income quintile, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quintile 1</td>
<td>191 (19.5)</td>
<td>2178 (21.8)</td>
<td>2585 (21)</td>
<td>24,750 (20.1)</td>
</tr>
<tr>
<td>Quintile 2</td>
<td>202 (19.6)</td>
<td>2071 (20.7)</td>
<td>2409 (19.7)</td>
<td>24,748 (20.1)</td>
</tr>
<tr>
<td>Quintile 3</td>
<td>205 (20.3)</td>
<td>2024 (20.2)</td>
<td>2455 (19.9)</td>
<td>24,561 (19.9)</td>
</tr>
<tr>
<td>Quintile 4</td>
<td>230 (22.9)</td>
<td>1993 (19.9)</td>
<td>2487 (20.1)</td>
<td>24,093 (19.5)</td>
</tr>
<tr>
<td>Quintile 5</td>
<td>173 (17.8)</td>
<td>1744 (17.4)</td>
<td>2395 (19.4)</td>
<td>25,158 (20.4)</td>
</tr>
<tr>
<td>Residence, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>887 (88.6)</td>
<td>9070 (90.6)</td>
<td>10,530 (85.4)</td>
<td>106,516 (86.4)</td>
</tr>
<tr>
<td>Rural</td>
<td>114 (11.4)</td>
<td>940 (9.4)</td>
<td>1801 (14.6)</td>
<td>16,794 (13.6)</td>
</tr>
<tr>
<td>Cancer diagnosis site, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rectum</td>
<td>418 (41.8)</td>
<td>N/A(^d)</td>
<td>3848 (31.2)</td>
<td>N/A</td>
</tr>
<tr>
<td>Left colon</td>
<td>410 (41)</td>
<td>N/A</td>
<td>5210 (42.3)</td>
<td>N/A</td>
</tr>
<tr>
<td>Right colon</td>
<td>102 (10.2)</td>
<td>N/A</td>
<td>2232 (18.1)</td>
<td>N/A</td>
</tr>
<tr>
<td>Transverse colon</td>
<td>55 (5.5)</td>
<td>N/A</td>
<td>753 (6.1)</td>
<td>N/A</td>
</tr>
<tr>
<td>Cancer diagnostic stage, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>270 (27)</td>
<td>N/A</td>
<td>2340 (19)</td>
<td>N/A</td>
</tr>
<tr>
<td>III</td>
<td>351 (35)</td>
<td>N/A</td>
<td>3644 (29.6)</td>
<td>N/A</td>
</tr>
<tr>
<td>II</td>
<td>205 (20.5)</td>
<td>N/A</td>
<td>2996 (24.3)</td>
<td>N/A</td>
</tr>
<tr>
<td>I</td>
<td>143 (14.3)</td>
<td>N/A</td>
<td>2680 (21.7)</td>
<td>N/A</td>
</tr>
<tr>
<td>0</td>
<td>32 (3.2)</td>
<td>N/A</td>
<td>671 (5.4)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

\(^a\)EAO-CRC: early age-onset colorectal cancer.
\(^b\)AAO-CRC: average age-onset colorectal cancer.
\(^c\)Cancer-free controls for individuals with AAO-CRC were not analyzed for study purposes but reported demographic characteristics for completeness.
\(^d\)N/A: not applicable.

We further assessed patterns of prescription medication use among EAO-CRC cases stratified by sex and stage. Multimedia Appendix 4 shows medication classes that represent ≥1% (n ≥130 prescriptions) of all prescription events in the year before EAO-CRC diagnosis according to sex. We observed a higher number of prescriptions (n=7295) representing 56.2% of all events among 420/551 (76.2%) male EAO-CRC cases. In contrast, 377/450 (83.8%) female EAO-CRC cases had a lower number of prescriptions (n=5694) representing 43.8% of events. In terms of frequency of prescriptions by sex, we found higher fills for antidepressants (n=1075, 14.7% male patients and n=623, 10.9% female patients), antiepileptics (n=711, 9.8% male patients and n=421, 7.4% female patients), gastrointestinal drugs (n=582, 8% male patients and n=213, 3.7% female patients), as well as pain-related medications such as opioids (n=426, 5.8% male patients and n=219, 3.9% female patients) and other analgesics and antipyretics (n=79, 1.1% male patients and n=33, <1% female patients) for male patients with EAO-CRC than female patients with EAO-CRC. When EAO-CRC cases were stratified by stage, we observed the following prescription events among individuals: stage I (1620 prescription events in 112, 78.3% cases), stage II (3523 prescription events in 167, 81.5% cases), stage III (3226 prescription events in 283, 80.6% cases), and stage IV (4620 prescription events in 209, 77.4% cases). As seen visually by the blue bars in Multimedia Appendix 5, drugs belonging to the nervous system class were the most represented across all 4 stages. Of note, when considering number of prescriptions, those among stage IV EAO-CRC cases represented 35.6% (n=4620) of all prescription events in contrast to those among stage I EAO-CRC cases, which represented 12.5% (n=1620) of all prescription events. Antidepressants were the most filled medications among individuals diagnosed at stage II (n=188, 14.2%) and stage IV (n=880, 19%; Multimedia Appendix 5). Of interest, gastrointestinal drugs were the most used in stage IV EAO-CRC cases (n=510, 11%). Topical agents for the treatment of hemorrhoids and anal fissures were mostly filled by stage III EAO-CRC cases (n=128, 4%). Drugs for constipation were the...
highest used in stage II EAO-CRC cases (n=92, 2.6%) and lowest in stage I EAO-CRC cases (n=22, 1.4%).

For further context, when we analyzed 12,331 AAO-CRC cases, we observed a total of 317,271 prescription events among 10,979 (89%) individuals (Multimedia Appendix 6), mean of 28.9 (SD 83.9) prescriptions (median 13.0) per AAO-CRC case. While antidipressants (n=1698, 13.1%) and antiepileptics (n=1132, 8.7%) were the top 2 most frequently filled medications among EAO-CRC cases, these drug classes were observed to be the third and seventh most used medications among AAO-CRC cases (n=15,097, 4.8% and n=10,689, 3.4%, respectively). Instead, the AAO-CRC group showed lipid modifying agents (n=21,898, 6.9%) and angiotensin-converting enzyme inhibitors (n=16,292, 5.1%) as the top 2 most used medication classes. Drugs that may be used to treat symptoms associated with potential CRC diagnosis were more frequently filled among EAO-CRC cases than AAO-CRC cases including gastrointestinal drugs (EAO-CRC: n=795, 6.1% and AAO-CRC: n=14,964, 4.7%), nonsteroidal anti-inflammatory and antirheumatic products (EAO-CRC: n=449, 3.5% and AAO-CRC: n=3430, 1.1%), topical agents for treatment of hemorrhoids and anal fissures (EAO-CRC: n=275, 2.1% and AAO-CRC: n=1672, <1%), and drugs for constipation (EAO-CRC: n=250, 1.9% and n=2897, <1%). EAO-CRC cases also revealed a higher use of opioids (EAO-CRC: n=645, 5% and AAO-CRC: n=9602, 3%).

Discussion

Overview

Using population-based administrative data, we assessed patterns of prescription medications in the year before diagnosis among individuals with EAO-CRC to understand the role of medications in the pathway to diagnosis in a condition that has seen a considerable increase in incidence [2-4]. Among 1001 EAO-CRC cases, 12,989 prescriptions were filled in the year before diagnosis by 797 (79.7%) individuals. With respect to medications, antidipressants were most commonly filled (n=1698, 13.1%), followed by antiepileptics (n=1132, 8.7%) and gastrointestinal drugs (ie, drugs for peptic ulcer disease and gastroesophageal reflux disease; n=795, 6.1%). Sex-based analyses revealed that male EAO-CRC cases had a higher number of prescriptions (n=7295, 56.2% of prescription events) but at a lower proportion (420/551, 76.2%), whereas female EAO-CRC cases had a lower number of prescriptions (n=5694, 43.8% of prescription events) but at a higher proportion (377/450, 83.8%).

Principal Findings and Comparison to Prior Work

Given the increasing risk of EAO-CRC [4] and reported diagnostic delays in prior studies [19,20], we were particularly interested in studying the patterns of prescription medication use leading to diagnosis in individuals with EAO-CRC and understanding potential diagnostic opportunities. To our knowledge, this study is the first to assess patterns of prescription use before diagnosis of EAO-CRC. In 2017, using Danish nationwide health registries on cancer and prescription drugs, Pottegård and Hallas [6] assessed the new use of prescription drugs among patients with lung, breast, colon, and prostate cancers 24 months preceding their cancer diagnosis. Authors found similar patterns of drug use between cancer cases and population controls in the 24- to 12-month period before cancer diagnosis. Among colon cancer cases, authors showed an increase in the use of prespecified drug classes that were likely prescribed for symptoms relating to their cancer, namely, proton pump inhibitors, laxatives or drugs against diarrhea, and opioid analgesics. However, this study did not characterize participants, and as such, it is not feasible to draw findings according to age as well as sex and stage, as with our study.

With respect to prescription medication use specifically among individuals with CRC, a 2021 cohort study by Engeland et al [21] using data from the Cancer Registry of Norway primarily assessed prescription medications after diagnosis but also reported on use in the year before diagnosis. Authors evaluated a prespecified list of drugs according to 5 major categories and reported the top 3 most commonly used drug groups in the year before diagnosis such as those for cardiovascular diseases (use prevalence 24.8%); endocrine, nutritional, and metabolic diseases (use prevalence 17.8%); and mental and behavioral disorders (use prevalence 6.7%). Although the study included patients with CRC aged 20-84 years, there was no reporting of drug use according to age groups. Furthermore, with 530 individuals in the 20- to 39-year age category comprising 2% of the study population, reported findings largely reflect drug use among older patients with CRC.

Indeed, this study provides a better understanding of patterns of prescription medication use specifically in EAO-CRC. In contrast to the aforementioned studies [6,21], which assessed prespecified lists of drugs based on reimbursement, we were able to assess all prescriptions, regardless of payer, given comprehensive capture in the PharmaNet database. At the outset, we initially assumed that the most common prescriptions filled during the year of diagnosis were for gastrointestinal and pain-based on previously reported symptoms of EAO-CRC [22]. Indeed, among the top 10 classes of most frequently filled prescriptions by EAO-CRC cases were gastrointestinal drugs for peptic ulcer disease and gastroesophageal reflux disease (third), opioids (fourth), anti-inflammatory and antirheumatic drugs and nonsteroids (sixth), topical agents for hemorrhoids and anal fissures (ninth), and drugs for constipation (10th). We believe the increased use of these drugs for EAO-CRC symptoms in the year prior to diagnosis may be the early manifestations of red flag signs and symptoms of CRC. A 2023 population-based case-control study by Fritz et al [23] identified 4 red-flag signs and symptoms (rectal bleeding, abdominal pain, diarrhea, and iron-deficiency anemia) that were associated with a heightened risk of EAO-CRC between 3 months to 2 years preceding diagnosis (ORs range between 1.34 and 5.13). These red flag symptoms align with the clinical indications of our results, where gastrointestinal drugs, pain medications, and rectal medications were among the top 10 classes of most frequently filled prescriptions by EAO-CRC cases in the year prior to diagnosis. These results highlight the importance of ensuring individuals younger than 50 years consistently presenting with these early warning signs, and symptoms or medication use patterns are being given ample opportunities for further work-up and early detection of CRC at their health care interactions. Stratified analyses by sex and stage further
reveal patterns such as higher use of pain-related medications and gastrointestinal drugs by male EAO-CRC cases. Our findings also suggest sex differences in health care use in terms of more frequent prescriptions among a smaller number of male EAO-CRC cases compared to less frequent prescriptions among a greater number of female EAO-CRC cases. With respect to stage, gastrointestinal drugs were most used in stage IV EAO-CRC cases, topical agents for treatment of hemorrhoids and anal fissures were by stage III EAO-CRC cases, and drugs for constipation were the highest used in stage II EAO-CRC cases and lowest in stage IV EAO-CRC cases. In contextualizing findings with those of controls, while opioids (fourth), gastrointestinal drugs (fifth), and nonsteroidal anti-inflammatory and antirheumatic drugs (seventh) were among the top 10 classes of filled prescriptions by matched cancer-free controls, they were at a lower frequency than EAO-CRC cases. Interestingly, topical agents for hemorrhoids and anal fissures and drugs for constipation were not among ≥1% (n=1748 prescriptions) of prescription events among controls.

Our findings on patterns of prescription medication use before diagnosis support a study rationale of exploring targets for raised awareness and education on the increasing risk of EAO-CRC to allied health care providers, particularly pharmacists. With patients reportedly seeing pharmacists 1.5 to 10 times more frequently than primary care physicians [24], these may represent windows of opportunity for education or identification of risks for diseases, including cancer. A survey of community pharmacists suggests that patients have long sought advice from pharmacists about possible cancer signs and symptoms [25]. With respect to CRC, pharmacists are gaining recognition for their roles in the initiation of average age screening in various jurisdictions [26-28]. In the United States, a 2-phased study showed high satisfaction among individuals from limited-income populations with pharmacists speaking to them regarding CRC screening [27]. In Spain, evaluation of a population-based CRC screening program showed high adherence by participating pharmacies (82.4%) with respect to distributing fecal immunochemical test kits and a high return rate by invitees (93.5%), demonstrating the important role that pharmacists play in the program [29]. There is indeed potential to expand on pharmacists’ roles when it comes to educating individuals regarding CRC, including younger adults about EAO-CRC. To date, calls to action have largely focused on increasing awareness among primary care physicians on the increasing risk of EAO-CRC [30,31]; however, it is also important to consider other health care providers, particularly pharmacists, given their accessibility and as prescriptions represent a frequent health care encounter prior to CRC diagnosis.

Aside from patterns of prescription medication use, a noteworthy finding from this study is that antidepressants represent the top prescribed drug class for EAO-CRC cases in the year before diagnosis, representing 13.1% (n=1698) of all prescription events. For context, antidepressants were also the top prescribed drug class for matched cancer-free controls but at a lower frequency, 9.9% (n=17,262). For further context, among AAO-CRC cases, antidepressants were the third most prescribed drug class (n=15,097, 4.8%) after angiotensin-converting enzyme inhibitors (n=16,292, 5.1%) and lipid modifying agents (n=21,898, 6.9%). A potential reason for this finding is a diagnostic delay of CRC that commonly occurs in the young patient population [20], which may lead to anxiety and depressive symptoms [32]. A systematic review that compared the delays and outcomes between younger and older patients with CRC found that younger patients are at a higher risk of experiencing delays from symptom onset to presentation, as they are not eligible for screening [20]. Consequently, a delay in cancer diagnosis in the younger population is associated with an increased risk of anxiety and depression [32]. A cross-sectional study in 2022 found that patient intervals (symptom onset to first seeing a general practitioner) of ≥1 month were associated with greater depression (aOR 1.7, 95% CI 1.1-2.5) compared to <1 month and having ≥3 prereferral general practitioner consultations were associated with greater anxiety (aOR 1.6, 95% CI 1.1-2.3) compared to 1-2 consultations [32]. The main reasons that could contribute to the increased risk of emotional distress in the adolescents and young adult population prior to a diagnosis include patients’ persistent symptoms being dismissed due to young age, unresolved symptoms, and the fear of a potential cancerous diagnosis [32,33]. Furthermore, a 2022 cohort study that used the same administrative databases as this study found that compared to individuals without cancer, those with EAO-CRC did not have a higher onset of depression after diagnosis (adjusted hazard ratio [aHR] 1.00, 95% CI 0.92-1.10) [34]. However, individuals with EAO-CRC had a 41% higher risk of onset of depression after diagnosis compared to individuals with AAO-CRC (aHR 1.41, 95% CI 1.25-1.60) [34]. Since we were not able to link indications to prescription events, we do not know whether antidepressants were prescribed for depression or for other reasons, such as pain. Nonetheless, findings in this study suggest a substantial burden of depression even before EAO-CRC diagnosis, which further indicates the need for person-centered mental health services for individuals with EAO-CRC across the entire spectrum of care.

**Strengths and Limitations**

The strengths and limitations of this study warrant discussion. We drew EAO-CRC cases and controls from population-based administrative health databases, namely Population Data BC and the BC Cancer Registry, which capture data on approximately 95% of all cancer cases in the province [9]. The BC Cancer Registry is reviewed annually for quality, completeness, and accuracy by the North American Association of Central Cancer Registries [9]. Nevertheless, this study is vulnerable to inherent limitations with administrative health data, which are not collected for research purposes. Although we have data on cancer stage, it is important to note this information in the BC Cancer Registry is not acquired using a systematic approach with sources including death certificates, pathology reports, and death certificates. Finally, administrative databases in BC do not yet capture information on the social construct of gender, and as such, we are not able to incorporate this into our analysis.

**Conclusions and Future Directions**

Altogether, using generalizable, population-based data, including a complete capture of all prescriptions, we delineated patterns...
of medication use before diagnosis of EAO-CRC. Our findings suggest a high frequency of prescription fills in the year before diagnosis of EAO-CRC, including for drugs to treat commonly reported symptoms of EAO-CRC. As efforts continue to raise awareness on the increasing risk of EAO-CRC, our findings provide support for also considering the role of other health care providers, particularly pharmacists. Altogether, prescription medications represent a common and potentially, frequent, point-of-contact with the health care system and thus may lend to a better understanding of trajectories for individuals with EAO-CRC.

Acknowledgments
This research was funded by a project grant from the Canadian Institutes of Health Research, “Examining the epidemiology, treatment, and outcomes in young-onset colorectal cancer” (funding reference PJT-159467). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. MADV holds a Tier 2 Canada Research Chair.

Data Availability
The data that support the findings of this study are available from Population Data BC but restrictions apply to the availability of these data, which were used under license for this study, and so are not publicly available. Data are available from Population Data BC through a data access request. Therefore, the data sets generated and analyzed during this study are not publicly available due to strict data sharing agreements with the BC Ministry of Health but are available from the corresponding author on reasonable request.

Authors’ Contributions
Vienna C contributed to conceptualization, formal analysis, investigation, methodology, project administration, visualization, data interpretation, and writing original draft. ECS contributed to data curation, formal analysis, investigation, methodology, software, validation, and visualization. Vicki C and RG contributed to conceptualization, investigation, visualization, data interpretation, and writing original draft. MADV contributed to funding acquisition, conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, visualization, data interpretation, and writing original draft. All authors reviewed and edited the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Overview of data sources and study sample (dashed arrow indicates linkages between databases using personal health numbers, which are then deidentified or scrambled).

[PDF File (Adobe PDF File), 90 KB - cancer_v10i1e50402_app1.pdf]

Multimedia Appendix 2
Anatomical Therapeutic Chemical Level 1 groups.

[DOCX File, 14 KB - cancer_v10i1e50402_app2.docx]

Multimedia Appendix 3
Frequency of prescriptions in the year before diagnosis for early age-onset colorectal cancer cases and cancer-free controls according to Anatomical Therapeutic Chemical Level 3 Classification.

[DOCX File, 21 KB - cancer_v10i1e50402_app3.docx]

Multimedia Appendix 4
Frequency of prescriptions in the year before diagnosis among male and female early age-onset colorectal cancer cases, according to Anatomical Therapeutic Chemical Level 3 Classification.

[DOCX File, 21 KB - cancer_v10i1e50402_app4.docx]

Multimedia Appendix 5
Bar charts showing percentage of prescriptions for the top 10 drug classes by Anatomical Therapeutic Chemical Classification Level 3 code, according to stage for early age-onset colorectal cancer cases.

[PDF File (Adobe PDF File), 423 KB - cancer_v10i1e50402_app5.pdf]
Multimedia Appendix 6

Frequency of prescriptions in the year before diagnosis for early age-onset colorectal cancer and average age-onset colorectal cancer cases according to Anatomical Therapeutic Chemical Level 3 Classification.

[DOCX File, 21 KB - cancer_v10i1e50402_app6.docx ]

References


Abbreviations

AAO-CRC: average age-onset colorectal cancer
aHR: adjusted hazard ratio
aOR: adjusted odds ratio
ATC: Anatomical Therapeutic Chemical
BC: British Columbia
CRC: colorectal cancer
EAO-CRC: early age-onset colorectal cancer
ICD-O-3: International Classification of Diseases for Oncology, Third Edition
OR: odds ratio

Please cite as:
Patterns of Prescription Medication Use Before Diagnosis of Early Age-Onset Colorectal Cancer: Population-Based Descriptive Study
JMIR Cancer 2024;10:e50402
URL: https://cancer.jmir.org/2024/1/e50402
doi:10.2196/50402
PMID:
bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.
Viewpoint

“Notification! You May Have Cancer.” Could Smartphones and Wearables Help Detect Cancer Early?

Suzanne E Scott1, BSc, MSc, PhD; Matthew J Thompson2, BSc, MBChB, MPH, DPhil

1Centre for Cancer Screening, Prevention and Early Diagnosis, Wolfson Institute of Population Health, Queen Mary University of London, London, United Kingdom
2Department of Family Medicine, University of Washington, Seattle, WA, United States

Corresponding Author:
Suzanne E Scott, BSc, MSc, PhD
Centre for Cancer Screening, Prevention and Early Diagnosis
Wolfson Institute of Population Health
Queen Mary University of London
Charterhouse Square
London, EC1M 6BQ
United Kingdom
Phone: 44 2078823550
Email: suzanne.scott@qmul.ac.uk

Abstract

This viewpoint paper considers the authors’ perspectives on the potential role of smartphones, wearables, and other technologies in the diagnosis of cancer. We believe that these technologies could be valuable additions in the pursuit of early cancer diagnosis, as they offer solutions to the timely detection of signals or symptoms and monitoring of subtle changes in behavior that may otherwise be missed. In addition to signal detection, technologies could assist symptom interpretation and guide and facilitate access to health care. This paper aims to provide an overview of the scientific rationale as to why these technologies could be valuable for early cancer detection, as well as outline the next steps for research and development to drive investigation into the potential for smartphones and wearables in this context and optimize implementation. We draw attention to potential barriers to successful implementation, including the difficulty of the development of signals and sensors with sufficient utility and accuracy through robust research with the target group. There are regulatory challenges; the potential for innovations to exacerbate inequalities; and questions surrounding acceptability, uptake, and correct use by the intended target group and health care practitioners. Finally, there is potential for unintended consequences on individuals and health care services including unnecessary anxiety, increased symptom burden, overinvestigation, and inappropriate use of health care resources.

(JMIR Cancer 2024;10:e52577) doi:10.2196/52577

KEYWORDS
wearables; early diagnosis; cancer; challenges; diagnosis; wearable; detect; detection; smartphone; cancer diagnosis; symptoms; monitoring; monitor; implementation; anxiety; health care service; mobile phone

Introduction

There is growing use of smartphones, wearables, and other technologies in health and wellness, either as consumer products or medical devices. The National Health Service (NHS) Long Term Plan [1] anticipates that in 10 years, people will have “the option for their physiology to be effortlessly monitored by wearable devices. People will be helped to stay well, to recognize important symptoms early, and to manage their own health, guided by digital tools.” Similarly, in 2020, the US Food and Drug Administration (FDA) launched a Digital Health Innovation Action Plan to encourage digital health innovation as “digital health technologies can empower consumers to make better-informed decisions about their own health and provide new options for facilitating prevention, early diagnosis of life-threatening diseases, and management of chronic conditions outside of traditional care settings” [2]. Wearables are devices that can be worn to detect and monitor biometric data such as heart rate, blood oxygen saturation, sleep pattern, or temperature while the wearer continues their normal routines. A further category of wearables involves skin patches used to measure biochemical signals (ie, glucose) on a continuous basis that are increasingly being considered as a standard of care for individuals with certain conditions (eg, diabetes) [3]. While most wearables have been wrist-worn devices, similar physiologic signals are now being generated from other devices.
such as rings or earbuds. Smartphones are also increasingly being used for health and wellness; they have the advantage of far higher use (compared with wearables), and a growing array of different sensors are routinely embedded. Could smartphones and wearables help detect cancer and, importantly, detect cancer earlier in its disease course when it is more likely to be localized and with a better prognosis? This paper provides an overview of the scientific rationale as to why these technologies could be valuable for early detection of cancer, the potential barriers to successful implementation, and the next steps for research and development.

**Potential of Smartphones and Wearables for Early Detection of Cancer**

While national cancer screening programs offer the opportunity to detect cancer or precancerous lesions in asymptomatic individuals, routine screening currently only accounts for the minority (<10%) of cancer diagnoses [4,5]. The predominant route to a cancer diagnosis is symptomatic presentation to ambulatory care. Thus, the diagnosis of cancer heavily relies on patients’ ability to notice and attend to relevant bodily changes and their decision to consult a health care professional [6,7]. However, noticing relevant bodily changes is challenging given the multiple subtle changes that may signal cancer among the plethora of daily bodily changes; fluctuations of normal bodily processes; self-limiting, transient symptoms; and the presence of chronic conditions. The signal-to-noise ratio is weak. This issue is exacerbated by individuals’ limited ability to accurately interpret vague bodily changes, many of which can be associated with cancer (eg, fatigue, weight loss, and stomach upset). This is because our awareness, attention, and interpretation are affected by expectations; emotions; beliefs; and biological, environmental, sociodemographic, and contextual factors [8-11]. Furthermore, symptoms may evolve very slowly over time, making it difficult to notice subtle changes. It is reported that the predominant risk factor for delay in seeking help following the detection of cancer symptoms is the "lack of interpretation by patients of the serious nature of their symptoms" [12].

Smartphones and wearable technologies have the potential to facilitate the detection and tracking of bodily changes that might otherwise be dismissed or interpreted as only needing self-medication rather than the attention of a health care professional. There is emerging data about early, subtle signs of cancer, and some of these may be amenable to detection by electronic sensors and monitoring of behavior (see Table 1).

**Table 1.** Potential signals of cancer that can be measured using sensors in smartphones or wearables.

<table>
<thead>
<tr>
<th>Sensors currently available on some smartphones or wearables</th>
<th>Examples of signals for health features that could be related to cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio signals from microphones [13]</td>
<td>• Changes in cough and breathing difficulty (associated with lung cancer) [14]</td>
</tr>
<tr>
<td>• Changes in voice such as hoarseness (associated with head and neck cancer and lung cancer) [14]</td>
<td></td>
</tr>
<tr>
<td>GPS location tracking and activity tracking [15]</td>
<td>• Reduced activity resulting from fatigue (associated with multiple cancers) [14]</td>
</tr>
<tr>
<td>Image capture and analysis [16-18]</td>
<td>• Anemia detected from images of the skin or eyes (associated with multiple cancers) [19]</td>
</tr>
<tr>
<td>• Jaundice detected from images of the skin or eyes (associated with pancreatic cancer) [20]</td>
<td></td>
</tr>
<tr>
<td>• Changes in skin lesions (associated with skin cancer) [14]</td>
<td></td>
</tr>
<tr>
<td>Temperature measurement [21]</td>
<td>• Rise in temperature (associated with pancreatic cancer) [22]</td>
</tr>
<tr>
<td>Body composition measurement using image analysis and electro dermal activity [23]</td>
<td>• Weight loss (associated with multiple cancers) [14]</td>
</tr>
<tr>
<td>Photoplethysmogram [24]</td>
<td>• Anemia (associated with multiple cancers) [19]</td>
</tr>
</tbody>
</table>

Sensors could allow the detection of changes prior to them being noticed or interpreted as symptoms, for example, a reduction in activity prior to fatigue or changes in food consumption prior to weight loss. There is recent evidence that monitoring day-to-day purchases could detect an increase in over-the-counter pain and indigestion medication 8 months prior to ovarian cancer diagnosis [25]. This demonstrates how tracking and monitoring change over time could allow insight into emerging disease. This is particularly useful for clinicians working in health care settings with limited time and resources and where cancer is a relatively rare occurrence among the burden of other diseases. In addition to the detection of signals, smartphones and wearables could alert the user to the need for health care consultation and provide an endorsement to seek care. This could overcome the commonly reported barrier to presentation (“concern about bothering the doctor”) that arises when there is uncertainty about the need for care [26-28].

**Potential Barriers to Successful Implementation**

**Overview**

Despite the promise of smartphones and wearables for early detection of cancer, there are several hurdles to implementation that require attention. Key barriers to success are outlined here, alongside suggestions for how these may be addressed with future research.
Signals and Sensors With Accuracy and Utility

The use and adoption of smartphones and wearables in this context require robust research into the selection of a signal, the development of sensors, and the generating evidence of accuracy and utility of those sensors in the real world. This includes identifying physiologic (or emerging pathophysiologic) signals that are most predictive of cancer, determining how often these need to be collected, and elucidating what other data would add precision to results (eg, age, risk factors, and presence of symptoms). For technology developers, sensors are usually designed and prioritized for a number of potential applications, mainly targeting overall health and wellness rather than diagnostic capabilities per se. Prioritizing these research and development efforts for cancer detection specifically over and above other priorities could be challenging to justify for business development reasons. Relatedly, the original intended commercial purpose of existing sensors may not have been connected to cancer detection. To make headway in this field of research, technology developers and device users will need to be willing to provide access to data for research. General Data Protection Regulation allows device users to share their data with third-party organizations under the right to portability. Developing systems to facilitate data sharing, in formats compatible with health data, could allow the generation of new data sets to signal cancer risk. This will prevent duplication of effort and maximize the use of existing data for public benefit.

While initial evidence on the accuracy of a sensor to detect a given health signal could involve case-control studies (eg, individuals recently diagnosed with cancer and matched controls), subsequent research would likely require large prospective cohorts. Further, given the weak signal-to-noise ratio, it is likely that signals from wearables or smartphones alone might lack sensitivity or specificity. Therefore, research that combines signals from wearables or smartphones with other digital sources of data (eg, symptoms recorded in health records and initial laboratory tests in primary care) will almost certainly be needed to demonstrate sufficient accuracy and utility in target populations. This was recently highlighted in a systematic review [29] of artificial intelligence technologies for skin cancer detection. Despite an abundance of digital products, the review highlighted that there has been very little testing in low-prevalence populations and limited data on the use of lower-quality images (eg, taken by patients or family physicians or using lower-quality phones), and as such, widespread adoption into practice has been limited [30].

Innovators also need to consider (and test) whether these new digital tools should and could detect more than 1 type of cancer (or detect other potentially important nonmalignant diseases; eg, cirrhosis in individuals with jaundice or depression in individuals with weight loss). Other considerations include who the target group is (eg, all adults or only those at higher risk of developing cancer), at what point in time (eg, certain age), and at what periodicity that group should begin using this technology for the detection of cancer. Further, it is well documented that symptom monitoring increases selective attention to the body, resulting in increased symptom reporting [31]; thus, the monitoring of symptoms could result in increased symptom burden. Development and testing need to determine the extent to which the monitoring of activity, symptoms, and other signals changes the outputs of those measurements [32,33].

Regulatory Challenges

Given the burden involved in fulfilling regulatory approvals for diagnostic devices, many smartphone technologies and wearables that could potentially have value for cancer detection will instead be introduced as products for overall health and wellness management. As the field expands, more guidance and standards for digital health tools are being introduced to ensure that they are not only safe and effective but also adoptable by the health care system [34]. For example, the recent UK National Institute of Health and Care Excellence (NICE) evidence standards framework [35] is intended to ensure new digital health technologies are clinically effective and offer value to the health care system. The framework includes standards concerning safety, quality, acceptability, bias mitigation, data practices, professional oversight, credibility with health professionals, safeguarding assurances, scalability, as well as evidence of real-world performance and use. In some countries, consumer protection regulations also determine standards that certain wellness features (eg, step counting and heart rate measurement) need to fulfill, even though these are not regulated medical devices. As specified in the CanTest framework for early cancer detection [36], research and development will benefit from this early specification of the criteria (eg, target product profiles) needed for successful digital products for cancer detection [37-39].

Ensuring Equity

A key issue of wearables and smartphone technologies is the potential for new innovations to exacerbate inequalities in cancer outcomes. Sociodemographic factors such as household income, age, level of education, and gender have been found to influence the use of mobile health (mHealth) technology [40-44] and there is “a real risk that the increased use of digital technologies will make care experiences and outcomes worse for some people (or communities)” [45]. Development of wearables and smartphone technologies for cancer detection should be conducted with an equity lens to focus on the views and needs of those living or working in more deprived areas and those at risk of lower health literacy (eg, those with lower educational level, older age, lower income, and ethnic minority groups) [46,47], so that cultural attitudes toward the use of technology, affordability, and access can be a focus in their development. Inclusion and diversity within the development and testing of sensors are vital so that products are not biased and work equally regardless of skin color or other physiological differences [29]. Affordability is also a crucial point. Even though smartphone use is extensive [48] in both higher- and lower-income countries, the availability and quality of sensors differ across brands and models of smartphones. Wearable devices have far lower penetration in most high-income countries and lower still in those individuals with lower socioeconomic status. If accurate and reliable sensors are only available on high-end devices, then the net result will be inequitable outcomes. The consideration of a reverse innovation approach may be useful here if it is possible, to focus testing on inexpensive, easy-to-use products that can be rolled out at scale.
Acceptability and Adherence

Crucial to the successful implementation of any innovation is early insight into the user perspective, including acceptability, uptake, and correct use by the intended target group [49-51]. Yet, this consultation is often omitted or occurs too late in mHealth implementation, resulting in user burden; technical issues; poor designs; and ultimately the lack of uptake, adherence, and impact of the technology [49,52-54]. Indeed, there is currently an absence of research on user perspectives on wearables and smartphone technologies for the detection of cancer. While key issues such as cost, motivation, comfort, ease of use, trust in data use, visibility, and interpretability of data are applicable across the spectrum of wearables and smartphone technologies in health [40,55], there may be additional, specific challenges for using these innovations for the detection of cancer.

In this context, acceptability also pertains to individuals’ willingness to share their data from smartphones and wearables with researchers, medical professionals, or private companies. Willingness to share data from wearables was reported to be lower than that for other commercial data [56]. Less than 15% (n=65) of survey respondents aged 60 years and older were willing to share wearable device data with academic research institutions and only 40% (n=423) of those aged 18-59 years were willing to do so. Trust in organizations and worry about data misuse have been shown to be a key factor in people’s willingness to share commercial data for health research [56]. Ensuring clear and transparent data use and data-sharing policies is vital for success. There are real concerns about the misuse of data, commercialization, and access to data by unauthorized people [57]. While data sharing is an essential component in the use of smartphones and wearables for cancer detection, data protection is equally as vital.

In research studies of wearables, dropout rates can be up to 44% [40], and nonadherence to wearing devices for the study duration can be up to 50% [58,59]. Nonadherence is likely to be even higher in people with preexisting comorbidities and for technologies requiring long-term engagement, as may be needed for cancer detection to track signals over time. Balancing the advantages and disadvantages of continuous versus intermittent measurements at certain intervals should be a key consideration.

It is, therefore, essential to investigate user perspectives in parallel with the development and potential future deployment of wearables and smartphone technologies for cancer detection. This also includes encompassing the views of clinicians who are involved in the ongoing surveillance and care of those with a history of cancer and would inevitably be involved in shared decision-making on the potential implementation of such technologies and, crucially, the ongoing clinical management of individuals whose sensors indicate signals of possible cancer. In general, primary care clinicians have not typically been deeply engaged in the implementation of other consumer-grade or regulated medical devices; understanding from these clinicians’ viewpoint on how they could use information from smartphones and wearables within their clinical care pathways is critical to any adoption [60].

Unintended Consequences on Individuals and Health Care Services

The exciting potential of wearable technologies for cancer detection must be considered alongside the possible negative consequences. As seen with other new developments in cancer detection, given the overall very low prevalence of cancer, even tests with very high specificity will lead to a large number of individuals with false positives. The subsequent need for investigation and resources needed to differentiate those with false versus true positives (ie, do have cancer) could be considerable. For the majority of individuals, this could lead to huge risks of overinvestigation and inappropriate use of health care resources [61]. For cancer detection specifically, we can anticipate a far higher potential for wearables and smartphone technologies to generate anxiety than for other conditions (eg, detection of sleep apnea, or detection of irregular heartbeat), especially among those already fearful of cancer recurrence. This is particularly relevant to the question of how “results” should be delivered to users, what support would be needed at that time, and whose responsibility this would be. On the other hand, wearable use may lead to a false sense of reassurance, leading to a lower perceived need to attend cancer screening or respond to symptoms (eg, “my wearable says I am healthy...there is no need to see my doctor”). This is similar to when a negative cancer screening test result can overly reassure patients and affect subsequent decisions to seek care [62]. These issues about the psychological and behavioral impact of smartphone technologies and wearables to detect cancer remain unexplored and need focused behavioral science research.

Conclusions

For most cancers, the time from detecting a bodily change to interpreting that change as requiring the advice of a health care professional constitutes a substantial proportion of the time prior to diagnosis. The detection of cancer remains one of the most prominent priorities of many health systems, governments, and private and public research funders [63,64], and “leaving no stone unturned” in technologies that could potentially improve early detection is a priority. The rapid advances in the hardware (ie, sensors) and software embedded in smartphones and wearables offer exciting and potentially untapped opportunities to detect early warning signs of cancer that may otherwise be missed. The research and development needed to advance this field include the selection of appropriate signals and development of effective sensors followed by robust clinical research into accuracy in real-world settings. This relies on the up-front specification of the target groups and their needs. Target product profiles should be developed specifically for cancer detection technologies, and innovators should consult these and consider regulatory challenges early in the process of development, to design products in line with the requirements of individuals, clinicians, and health care systems. The potential negative consequences of this type of technology should be acknowledged and investigated up-front, and mitigations should be incorporated into the design and implementation strategies. To avoid exacerbation of inequalities in cancer outcomes, research into the use of wearables and smartphone technologies in cancer detection should be done with an equity lens to ensure...
that products are developed for those who have poorer health outcomes, for whom new innovations could have the most impact. There is a need for research to explore the patient, public, and health care perspectives about the use of smartphones and wearables for the early detection of cancer while this field is in its infancy, so that these can be incorporated into product design to optimize acceptability and adherence, avoid unintended consequences, and maximize the chance of their success.

Acknowledgments
SES is supported by Barts Charity (G-001520; MRC&U0036). MJT received no funding in support of this manuscript.

Authors' Contributions
SES and MJT conceived the work that led to the submission. SES drafted the manuscript. SES and MJT revised the manuscript and approved the final version.

Conflicts of Interest
SES declares no conflicts of interest. MJT is an employee of Google and owns Alphabet stock. None of the opinions or views stated in this paper reflect the opinions or views of Google Inc.

References


Abbreviations

FDA: Food and Drug Administration
NICE: National Institute of Health and Care Excellence
NHS: National Health Service
mHealth: mobile health
Web-Based Scaffolds: The Feasibility of a Constructivist Approach to Oncology Fellow Learning

Sam Brondfield¹, MAEd, MD; Matthew Schwede², MD; Tyler P Johnson², MD; Shagun Arora¹, MD

¹Department of Medicine, University of California, San Francisco, San Francisco, CA, United States
²Department of Medicine, Stanford University, Palo Alto, CA, United States

Corresponding Author:
Sam Brondfield, MAEd, MD
Department of Medicine
University of California, San Francisco
505 Parnassus Ave
Room M1286, Box 1270
San Francisco, CA, 94143
United States
Phone: 1 4155144783
Email: sam.brondfield@ucsf.edu

Abstract
In this 2-institution feasibility pilot, oncology fellows used and updated freely available web-based learning tools (scaffolds) in a constructivist fashion.

(JMIR Cancer 2024;10:e52501) doi:10.2196/52501

KEYWORDS
constructivist learning; scaffolded learning; graduate medical education; fellowship training; oncology; feasibility; medical education; pilot study; study; online learning; online tool; online tools; remote learning; e-learning; training; cancer

Introduction
Succinct and updated oncology fellow learning materials are lacking. Additionally, fellow didactic learning often takes the form of passive lectures, which is undesirable [1,2]. Constructivist learning, wherein learners construct their own knowledge, is rare for fellows.

We piloted “scaffolds”—succinct slide sets shared across oncology trainees—and evaluated feasibility [3,4]. Throughout training, fellows can update the shared scaffolds in a constructivist fashion, thereby providing updated resources for themselves and colleagues.

Methods
Study Design
Two institutions participated—University of California, San Francisco (UCSF), and Stanford University. From 2018 to 2019, SB—a UCSF oncologist—designed 12 scaffolds, using Google Slides covering the solid tumor chapters from the American Society of Clinical Oncology’s Self-Evaluation Program (ASCO-SEP) textbook [5]. Hematology, gynecologic oncology, and neuro-oncology were omitted for this pilot. Scaffolds included text and images synthesized from ASCO-SEP and National Comprehensive Cancer Center guidelines. For brevity, the slides instructed fellows to adhere to length limits when making edits.

We emailed scaffold links to all first- to third-year UCSF (n=21) and Stanford University (n=27) oncology fellows in July 2019 and July 2020. Use was optional, and fellows could access and update the scaffolds anonymously at any time. Updates were audited by SB.

In December 2021, to evaluate feasibility outcomes (fidelity: degree to which the innovation was implemented as intended; appropriateness: perceived fit of the innovation; self-efficacy: belief in the ability to execute the innovation’s goals) [6], we reviewed updates tracked in Google Slides and conducted 2 voluntary feedback focus groups (UCSF: facilitated by SB; Stanford University: facilitated by MS—a Stanford University oncology fellow) with 4 fellows each. Focus group size was determined by responses to recruitment emails. Consent and demographic information were obtained. Participants did not need to use the scaffolds, as we were also exploring barriers to use. Focus groups were recorded and professionally transcribed. SB and MS independently reviewed the transcripts and generated themes through iterative discussion [7].
The scaffolds were updated in 2023 by SB (available on Google Drive) [8].

**Ethical Considerations**

UCSF and Stanford University institutional review boards granted exemption (#20-31645) and approval (#57766), respectively. Participants received an information sheet and verbally consented before each focus group. Transcripts omitted personal identifiers, and interviewers never revealed participant identities to the rest of the study team. Participants received a US $10 electronic gift card.

**Results**

**Fidelity**

From July 2019 to December 2021, fellows made 60 updates (Table 1), ranging from new trials to changes in management; none were erroneous. SB made 9 edits for brevity.

Table 1. Number of updates to solid oncology scaffolds during the pilot period (July 2019 to December 2021).

<table>
<thead>
<tr>
<th>Scaffolds</th>
<th>Updates by fellows (N=60), n</th>
<th>Updates by auditor (N=9), n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder/kidney/adrenal</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Breast</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Gastrointestinal (lower)</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Gastrointestinal (upper)</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Germ cell</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Head/neck</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Lung (nonsmall cell)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Lung (small cell/other thoracic)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Melanoma</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Prostate</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Salivary/thyroid</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>

** Appropriateness**

Focus group participants (N=8) were women and included Asian (n=3, 37.5%), White (n=3, 37.5%), Black (n=1, 12.5%), mixed-race (n=2, 25%), first-year (n=5, 62.5%), second-year (n=2, 25%), and third-year (n=1, 12.5%) fellows. Most (n=7, 87.5%) used the scaffolds. Qualitative analysis (Table 2) revealed that fellows felt the scaffolds were accessible and succinct learning tools, addressed the dearth of similar resources, served as effective preparation materials for clinical work and examinations, provided structured information for rapid reviews, and made interactions with complex resources easier.
Table 2. Qualitative analysis of transcripts from 2 oncology fellow focus groups (1 at the University of California, San Francisco, and 1 at Stanford University) that evaluated a pilot of solid oncology scaffolds (July 2019 to December 2021).

<table>
<thead>
<tr>
<th>Theme</th>
<th>Supportive quotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages</td>
<td>“[The scaffolds were] online and quickly accessible, for example on the shuttle on the way to work.”</td>
</tr>
<tr>
<td>Accessible, succinct resource</td>
<td>“There are few resources currently available for oncology fellows. [The scaffolds] filled a niche not currently filled by other resources.”</td>
</tr>
<tr>
<td>Addressed the dearth of similar resources</td>
<td>“[The scaffolds] were a security blanket…helpful for clinic prep and inpatient consults.”</td>
</tr>
<tr>
<td>Effective preparation materials for clinical work and examinations</td>
<td>“[The scaffolds] were helpful in that they provided frameworks…and approaches.”</td>
</tr>
<tr>
<td>Structured information for rapid reviews</td>
<td>“[The scaffolds] were helpful in that they provided frameworks…and approaches.”</td>
</tr>
<tr>
<td>Easier subsequent use of more complex resources</td>
<td>“The guidelines felt less ‘foreign’ after reviewing the scaffolds…[the scaffolds] helped with knowledge retention from more complex resources.”</td>
</tr>
<tr>
<td>Challenges</td>
<td>“I wasn’t sure whether my learning points were important enough to add to the scaffold.”</td>
</tr>
<tr>
<td>Lack of fellow confidence in updating the scaffolds</td>
<td>“I think fellows are probably less likely to update the scaffolds if they don’t feel responsible for them.”</td>
</tr>
<tr>
<td>Too simple and broad to help with nuanced patient care</td>
<td>“Clinical care is so nuanced…the scaffolds may be too broad to help with some clinical situations.”</td>
</tr>
<tr>
<td>Suggestions</td>
<td>“Maybe make them more visually appealing by including more figures or tables.”</td>
</tr>
<tr>
<td>Clarify purpose and the fact that scaffolds can be updated</td>
<td>“I would make it clear that the slides are editable and that fellows should update them.”</td>
</tr>
<tr>
<td>Facilitate opportunities for fellows to update scaffolds</td>
<td>“Asking fellows to update these might be good for their learning.”</td>
</tr>
</tbody>
</table>

Self-Efficacy
Qualitative analysis revealed barriers to updating the scaffolds—fellows’ lack of ownership over the scaffolds and low confidence regarding appropriate updates.

Discussion
Principal Results
This pilot explored the feasibility of implementing constructivist scaffolds for oncology fellows. We found evidence of fidelity and appropriateness and delineated next steps to optimize self-efficacy. The scaffolds [8] can be downloaded and modified to avoid generating institution-specific scaffolds from scratch. To promote ownership and confidence, we recommend assigning fellows to update the scaffolds under faculty mentorship.

Despite demonstrating superior outcomes when compared to passive lectures, constructivist learning is rarely studied at the fellowship level [9-11]. We recommend evaluating constructivist learning modalities, such as scaffolds, in graduate medical education to enhance learning outcomes.

Limitations
Though the focus groups suggested that multiple fellows used the scaffolds, Google Slides did not track how many fellows accessed or updated them. We did not incorporate multimedia components beyond images and tables (some needed to be removed before publication to respect copyright), nor did we include assessments in this pilot. We recommend that institutions consider incorporating multimedia content and assessments into the scaffolds. The number of focus group participants was small and not gender-diverse. Future studies should quantitatively evaluate usage patterns and user satisfaction to examine what factors drive utilization.

Conclusion
We piloted a novel constructivist approach to fellow learning and found evidence of feasibility. Oncology educators may use and modify the scaffolds [8] to jump-start constructivist education for fellows at their institutions. Educators in other fields may wish to apply this model to their specialties.
Acknowledgments
This study was funded by a University of California, San Francisco (UCSF), Academy of Medical Educators Education Innovations grant. The funder did not have a role in the study’s design, data collection, data analysis, data interpretation, manuscript writing, or decision to submit the manuscript for publication.

Data Availability
The data sets analyzed during this study are not publicly available due to institutional review board restrictions but are available from the corresponding author on reasonable request.

Authors’ Contributions
SB designed the scaffolds, conceived the study, conducted the quantitative analysis, and wrote the manuscript. SB and MS each conducted 1 focus group. SB and MS conducted the qualitative analysis. All authors contributed manuscript edits and approved the final manuscript for submission.

Conflicts of Interest
None declared.

References

Abbreviations
ASCO-SEP: American Society of Clinical Oncology’s Self-Evaluation Program
UCSF: University of California, San Francisco

https://cancer.jmir.org/2024/1/e52501 JMIR Cancer 2024 | vol. 10 | e52501 | p.390 (page number not for citation purposes)