

Original Paper

Leveraging Digital Twins for Stratification of Patients with Breast Cancer and Treatment Optimization in Geriatric Oncology: Multivariate Clustering Analysis

Pierre Heudel^{1*}, MSc, LLM, MD; Mashal Ahmed^{2*}, PhD; Felix Renard^{2*}, PhD; Arnaud Attye^{2*}, MD, PhD

¹Department of Medical Oncology, Centre Leon Bérard, Lyon, France

²GEODAISICS, Grenoble, France

*all authors contributed equally

Corresponding Author:

Pierre Heudel, MSc, LLM, MD
Department of Medical Oncology
Centre Leon Bérard
28 rue Laennec
Lyon, 69008
France
Phone: 33 478782952
Email: pierreetienne.heudel@lyon.unicancer.fr

Abstract

Background: Defining optimal adjuvant therapeutic strategies for older adult patients with breast cancer remains a challenge, given that this population is often overlooked and underserved in clinical research and decision-making tools.

Objectives: This study aimed to develop a prognostic and treatment guidance tool tailored to older adult patients using artificial intelligence (AI) and a combination of clinical and biological features.

Methods: A retrospective analysis was conducted on data from women aged 70+ years with HER2-negative early-stage breast cancer treated at the French Léon Bérard Cancer Center between 1997 and 2016. Manifold learning and machine learning algorithms were applied to uncover complex data relationships and develop predictive models. Predictors included age, BMI, comorbidities, hemoglobin levels, lymphocyte counts, hormone receptor status, Scarff-Bloom-Richardson grade, tumor size, and lymph node involvement. The dimension reduction technique PaCMAP was used to map patient profiles into a 3D space, allowing comparison with similar cases to estimate prognoses and potential treatment benefits.

Results: Out of 1229 initial patients, 793 were included after data refinement. The selected predictors demonstrated high predictive efficacy for 5-year mortality, with mean area under the curve scores of 0.81 for Random Forest Classification and 0.76 for Support Vector Classifier. The tool categorized patients into prognostic clusters and enabled the estimation of treatment outcomes, such as chemotherapy benefits. Unlike traditional models that focus on isolated factors, this AI-based approach integrates multiple clinical and biological features to generate a comprehensive biomedical profile.

Conclusions: This study introduces a novel AI-driven prognostic tool for older adult patients with breast cancer, enhancing treatment guidance by leveraging advanced machine learning techniques. The model provides a more nuanced understanding of disease dynamics and therapeutic strategies, emphasizing the importance of personalized oncology care.

JMIR Cancer 2025;11:e64000; doi: [10.2196/64000](https://doi.org/10.2196/64000)

Keywords: digital twins; artificial intelligence; breast cancer; older adult patients with cancer; treatment; geriatric oncology; geriatric; oncology; cancer; clustering analysis; therapeutic; older adult; elder; old; patients with cancer; decision-making tools; decision-making; manifold learning model; chemotherapy; comorbidities; comorbidity; health care

Introduction

Breast cancer is more commonly diagnosed in older populations, particularly among women aged 65 years and

older in wealthier countries. In the United States, the average age of breast cancer diagnosis is 62 years, and in 2020, women aged 70 years and older accounted for 30% of all new cases of the disease [1,2]. In the European Union, women

older than 65 years made up about 44% of all breast cancer cases [3]. However, treatment approaches for early-stage breast cancer in these older age groups are often inadequate and unclear, largely due to a lack of solid evidence and the unreliability of web-based tools for making decisions about additional therapies, leading to less than ideal treatment outcomes [4,5].

The treatment plan for breast cancer is tailored based on the cancer's characteristics, the patients' overall health status, and their personal preferences. Standard care for early-stage breast cancer usually involves surgery, and may also include radiation, as well as neoadjuvant or adjuvant systemic therapy, used alone or in various combinations. Crafting postsurgical treatment strategies for older patients with breast cancer is complex due to their typically compromised health and the lack of data from clinical trials, since older adults are seldom participants in such studies and are not well represented in meta-analyses that evaluate the effectiveness of adjuvant chemotherapy in reducing breast cancer mortality and improving survival rates [6,7]. Consequently, artificial intelligence (AI) has been investigated as a potential tool to support decision-making in the context of limited clinical trial evidence.

Early uses of AI in cancer treatment guidance involved knowledge-based systems [8,9]. Recently, a broader spectrum of machine learning methods has been examined to aid both clinicians and patients with breast cancer [10-15]. Nonetheless, most decision support tools are designed for patients aged between 18 and 65 years, reflecting the age group most studied, with limited research focusing on treatment outcomes for older patients with breast cancer [16-19]. The prognostic tool PREDICT [20], although popular, has shown limited effectiveness for older adult patients [21]. Adjutorium [22], which uses extensive datasets from the United Kingdom and the United States, provides more precise prognosis and treatment benefit predictions for breast cancer than PREDICT. Despite this, it primarily includes patients aged between 30 and 65 years, with fewer older patients in its datasets, and omits certain vital tumor information such as progesterone receptor (PR) status [19]. Another established tool, Adjuvant! Online, predicts 10-year overall survival, breast cancer survival, and recurrence rates, commonly used to inform expected outcomes from endocrine therapy and chemotherapy [23]. Its accuracy is questionable for older women with early-stage breast cancer, probably because it was trained on data with a maximum age limit of 69 years [24]. In a review by Engelhardt et al [25], various models could forecast breast cancer outcomes, typically based on genetic risk scores, but only Adjuvant! Online factored in comorbidity status. Yet, none had been thoroughly validated in older adult populations. The more recent PORTRET tool was designed to predict 5-year recurrence, overall mortality, and mortality from other causes in patients older than 65 years with early invasive breast cancer, as well as to estimate the benefits of adjuvant systemic treatment [26]. The tool's authors observed that their treatment effect estimates were based on data from pooled randomized clinical trials, which

might not be entirely applicable to older adults due to the typically selective nature of older participants in these trials.

This study aims to develop models that overcome the shortcomings of past research by using cohorts that accurately reflect the demographic of older patients with breast cancer and by leveraging a detailed dataset that includes administrative, biological, treatment, primary tumor, and survival information. Our latest research uses manifold learning, an advanced tool for nonlinear dimensionality reduction that excels in unraveling complex geometric relationships within high-dimensional data, revealing intricate connections between clinical factors.

We introduce a new prognostic and predictive tool tailored for older adult patients with breast cancer, providing postoperative treatment recommendations. This tool is distinctive in its consideration of the interdependencies among variables within a patient population. It acknowledges the relative importance of prognostic factors in a way that many existing models do not. Our findings are set to be extremely beneficial for oncologists when determining suitable adjuvant treatment approaches for older adult patients with breast cancer, taking into account the nuances of both tumor-related and patient-specific characteristics.

Methods

Recruitment

In this retrospective study, we examined pseudonymized data from women aged 70 years and older who received a diagnosis of early-stage breast cancer and underwent surgery with the intent to cure (either lumpectomy or mastectomy, with or without axillary lymph node dissection) at the French Léon Bérard Cancer Center from January 1997 to December 2016. The French Léon Bérard Cancer Center is a 300-bed comprehensive cancer center located in Lyon, France, serving more than 30,000 patients annually, with a multidisciplinary team of 2000 health care professionals and a catchment area covering southeast France.

The inclusion criteria were not limited by the breast cancer's histological or molecular characteristics, the size of the tumor, or the status of the lymph nodes. However, the study did exclude patients who had noninvasive in situ carcinoma without invasive carcinoma, HER2 (human epidermal growth factor receptor 2) positive breast carcinoma, or who presented with distant metastases at the time of surgery. HER2-positive breast cancer cases were excluded because these patients typically receive trastuzumab-based targeted therapies, which dramatically improved their prognosis following its widespread adoption for nonmetastatic breast cancer around 2005. In contrast, chemotherapy protocols for HER2-negative cases remained consistent during the treatment period of the patients included in this study, ensuring uniformity in therapeutic strategies and outcomes across the cohort. The research concentrated on the 5-year survival rates, selecting only those who had at least 5 years of follow-up and whose vital status information was available.

The database was constructed using ConSore, a data-mining application developed by UNICANCER [27]. The ConSore platform extracts data from the electronic health records of the Léon Bérard Cancer Center, integrating patient demographics, clinical variables, and treatment details. To ensure accuracy, each record was also subject to a manual verification process. Data compiled included demographic details and clinical features of patients at diagnosis, alongside comprehensive biological and disease-specific information, and the treatments administered.

We included the following characteristics for patients diagnosed with early-stage breast cancer: age; Eastern Cooperative Oncology Group performance status; BMI; comorbidities such as diabetes, heart failure, coronary artery disease, chronic obstructive pulmonary disease, and cognitive impairments; history of hospitalizations; and polypharmacy. We also gathered biological indicators at the time of diagnosis, which included hemoglobin levels, lymphocyte counts, and creatinine clearance. We extracted data on disease attributes including histological subtype, hormone receptor status, HER2 status, Scarff-Bloom-Richardson (SBR) grade, tumor count, size of the largest tumor, and the extent of lymph node involvement as per the Tumor, “Nodes,” “Metastases (TNM) classification [28]. The statuses of estrogen receptors (ERs), PRs, and HER2 were determined from the histopathological analysis of pretreatment biopsies. Hormone receptor negativity was classified when fewer than 10% of cells were stained for ER and PR. HER2 negativity was assigned when immunohistochemistry staining was below 1+. For tumors scoring 2+, further in situ hybridization tests were conducted to assess HER2 amplification [29]. Treatment data collected encompassed the type of surgery performed, lymph node dissection, and adjuvant treatments including radiotherapy, chemotherapy, and endocrine therapy.

Outcome, Predictors, and Predictive Power

Outcome was overall survival in 5 years. Due to the high percentage of missing values for cause of death, cancer-specific survival was not considered. Nine predictors were selected: age, tumor size (mm), tumor grade (defined as either SBR low: 1-2; or high: 3), number of affected ganglions, hormone-receptor status (positive if either estrogen or PRs were immunohistochemically present in $\geq 10\%$ of tumor cells; otherwise, patients were classified as triple negative), serum hemoglobin (g/dL) and lymphocyte count (G/L), BMI, and the presence of comorbidities.

The initial database, built using ConSore, compiled a range of clinical, biological, and disease-specific data, along with information on administered treatments. We aimed for a predictors representing a mixture of features typically tested before patients undergo treatment plans. Thus, we excluded features regarding treatments as (1) we wanted to gauge prediction accuracies based only on the initial testing of the patient, and (2) the efficacy of treatment strategies was also an outcome of interest in the study. We further excluded features with significant number of missing values so as to limit the loss of usable data. Creatinine was excluded due to

its high correlation with patient’s age and potential kidney disorders that are not uncommon in the study’s demographic. The feature was found to correlate with negative patient outcome, but this was independent of cancer and introduced a bias. Following these steps, 9 predictors were isolated, a list comprising both continuous and categorical variables, as well as an acceptable mixture of relevant biological and clinical features. Random Forest Classification (RFC) and Support Vector Classifier (SVC) were used to evaluate the predictive power of the selected features. We used 5-fold cross-validation to mitigate overfitting and ensure the validity of our results.

Model Development and Validation

Patients in the initial cohort with missing values for any of the 9 predictors were cut from the study. The remaining patients comprised the model development cohort. This was divided into reference and model data.

Reference Data and Digital Twins

The reference data inclusion criteria were positive outcome for survival in 5 years and remission without relapse by the last follow-up. The purpose of this group was to calibrate the our patented algorithm, generating digital twins for future test subjects. Digital twins refer to synthetic patient data derived from the reference group specifically similar in profile to a new test subject. The model uses these synthetic profiles to recognize complex variations within the test profile. Thus, digital twins are generated and used in the model to provide recommendations on a new patient but do not themselves constitute the result that a physician would need to interpret.

Model Data

The model data, distinct from the reference data to prevent data leaking, are the population that is run through the precalibrated model and scored against the reference group. The data are thus transformed from raw patient data to a numerical and standardized representation of their deviation from the reference group (their digital twins). The purpose of these transformed data is to populate the model with a range of patient profiles that will serve for future prognostic analysis.

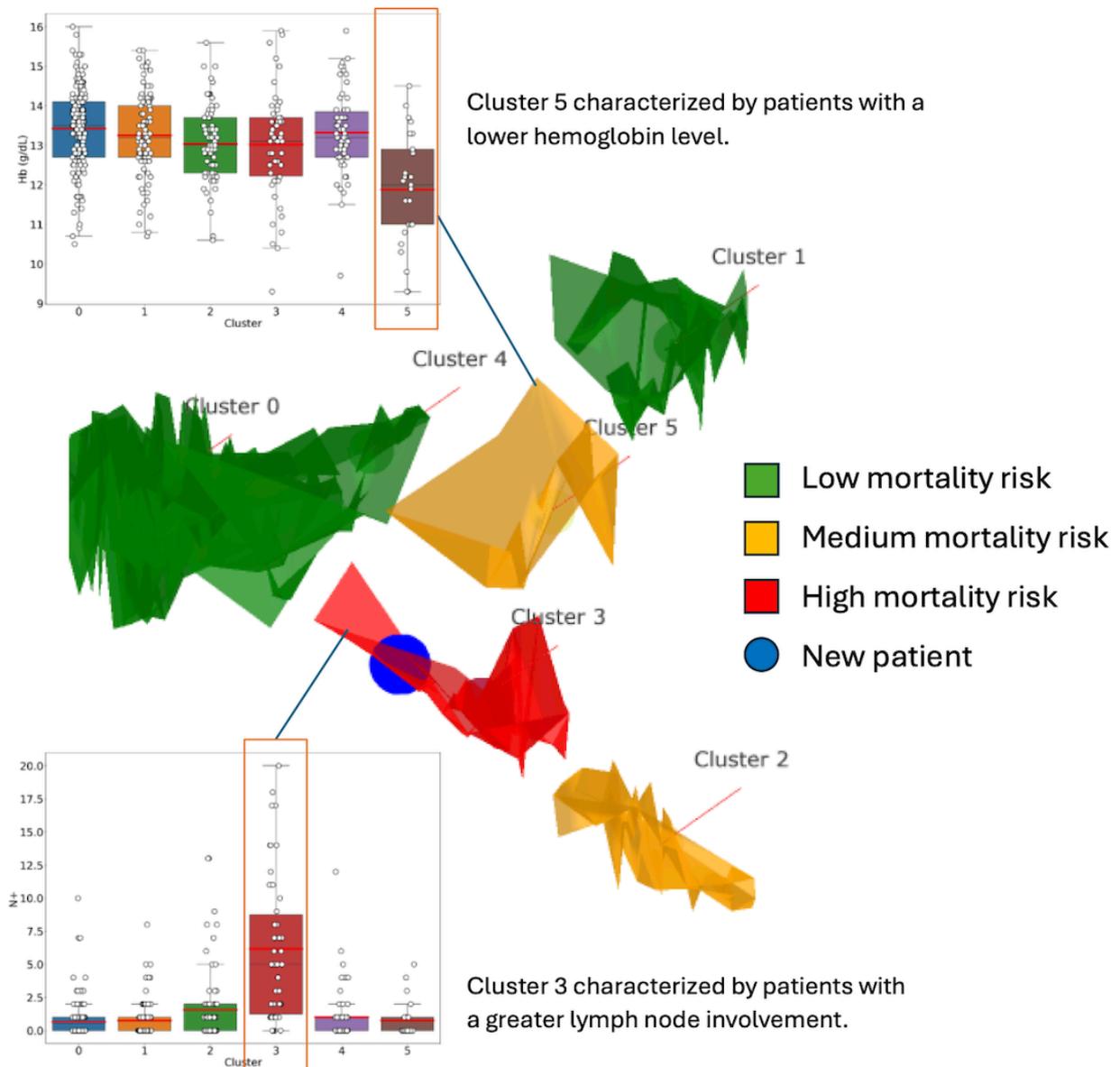
PaCMAP, Mean-Shift Clustering, and Manifold Visualization

The transformed model data underwent dimensionality reduction using PaCMAP (Pairwise Controlled Manifold Approximation) [30] to generate 3D data referred to as a manifold, permitting easy visualization. The data were then stratified using mean-shift clustering [31], a nonparametric, density-based clustering algorithm that can be used to identify clusters in a dataset (Figure 1). Each cluster represents a local group of similar patients in the 3D space. Clusters represent typical patient profiles in the overall population. The advantage of clustering is that it captures the variability of subjects of a subgroup for easy analysis. A better understanding of the cluster and its variability allows clinicians to assess whether a new test subject aligns well with the cluster

and to identify potential differences. When considering a new patient, estimates of prognosis and expected benefits of adjuvant treatment are ascertained by the examination of

cluster-specific treatment outcomes pertinent to the patient’s clinical profile.

Figure 1. Graphical representation of the 6 clusters of patients in the 3D manifold space. Patients in the reduced 3D space, or manifold, were grouped into clusters by their spatial distribution and profile similarity. Clusters were then colored based on the overall mortality rate of included patients. A newly tested patient is localized on the manifold and represented by a blue sphere.



Prediction of Chemotherapy Benefit

To estimate the benefit of chemotherapy, the position of a new patient is identified within the 3D manifold. Using the K-nearest neighbors algorithm, the 15 closest chemotherapy-treated patients and the 15 nearest non-chemotherapy-treated patients are pinpointed. Kaplan-Meier (KM) survival curves were plotted for each of these patient groups, providing a visual estimation of chemotherapy benefit for a clinical profile.

Validation of Treatment Benefit Predictions With Kullback-Leibler Divergence

To validate that the distributions of the 2 treatment subgroups are comparable, we used the Probability Density Function, which describes the spread of the data points in the 3D space. To measure the difference between these distributions, we applied the symmetrized Kullback-Leibler (KL) divergence, a statistical method that quantifies how much one distribution differs from another. To assess whether the observed difference was meaningful or just due to random chance,

we conducted a permutation analysis. This technique works by randomly shuffling the data multiple times to create many new random comparisons; comparing the real result with the random results allows us to determine whether the observed difference between the distributions was statistically significant. If distributions of 2 different treatment groups were found to be similar, they could be compared to provide a prediction of treatment benefit.

Model Stability Validation

The original model data were split into 2 groups: 70% (327) of every cluster was pooled into the training group, and the remaining 30% (139) was pooled into the test group. A new manifold learning process was applied to the training group, and the test group was then projected onto this newly generated manifold. Patients in the test and model groups from the same cluster of origin were compared to evaluate whether data points would exhibit similar distributions (appear in proximity to each other) in the new manifold space across 10 different manifold initializations.

Statistical Analysis

Kullback-Leibler Divergence and Permutation Test

The symmetrical KL divergence was used to measure the difference between 2 probability distributions. A permutation test was subsequently conducted to assess the significance of the observed KL divergence. This involved calculating the KL divergence for a large number of permutations of the combined datasets and comparing these values with the original KL divergence. The P value is calculated as the proportion of permutations where the KL divergence is as extreme as, or more extreme than, the original KL divergence calculated between the actual groups, thus providing a measure of how likely it was to observe a divergence as extreme as the original, under the null hypothesis of no difference between the distributions. Mathematically, this P value is the ratio of the number of permuted KL divergences that are equal to the original KL divergence or greater to the total number of permutations. A low P -value suggests that the observed difference in distributions is unlikely to have

occurred by chance, thus indicating a significant divergence between the 2 groups.

Survival Analysis using the KM Estimator and Log-Rank Test

The KM estimator was used to generate survival curves for different treatment subgroups. The log-rank test, a nonparametric test, was applied to compare the survival distributions and a P value was calculated to determine the statistical significance of the differences observed between the groups. A low P value suggests that the observed survival curves are significantly different. The statistical package used for the analysis is Lifelines 0.30.0 (Lifelines Developers) [32].

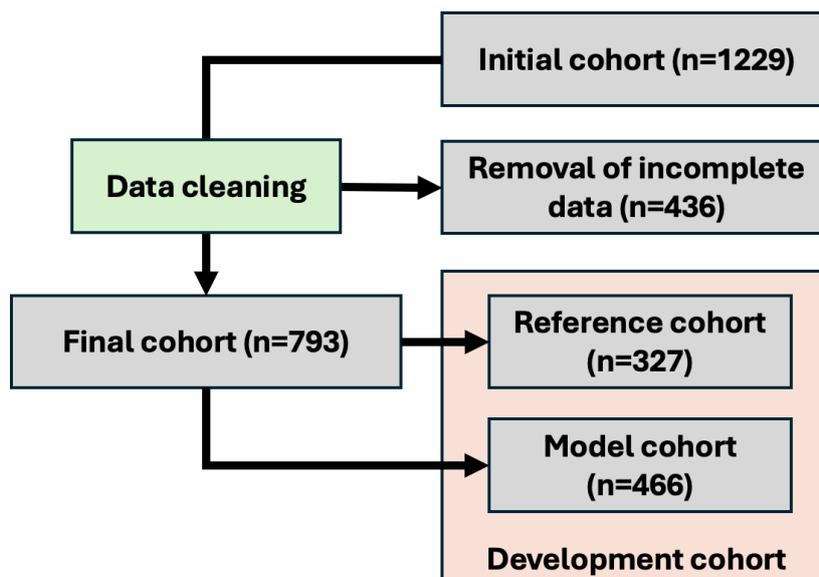
Ethical Considerations

This retrospective study involving human subjects was reviewed and approved by the French data protection authority, the Commission Nationale de l'Informatique et des Libertés, under authorization number 9191415, dated October 10, 2019. According to institutional and national guidelines, no additional approval from a research ethics board was required, as the data used were previously collected for clinical purposes. No new informed consent was required for this study. The analysis was conducted using data for which participants had provided general consent at the time of data collection. All data were pseudonymized prior to analysis to protect patient confidentiality. No identifiable personal information was retained in the research dataset. No compensation was provided to participants.

Results

Cohort Characteristics

A total of 1229 patients comprised the initial cohort. Of these, 793 (65%) remained after entries with missing values were removed (Figure 2). Eliminating the risk of introducing a bias, the initial cohorts' demographic and clinical characteristics were found to be strictly similar to that of the final cohort and are summarized in Tables 1 and 2.

Figure 2. Flowchart of data construction.**Table 1.** Patient characteristics of the initial cohort (N=1229).

Characteristics	Participants
Age at diagnosis (years), n (%)	
70-74	580 (47)
75-79	331 (27)
80-84	204 (17)
85-89	93 (8)
>90	20 (2)
Performance status, n (%)	
0	339 (28)
1	322 (26)
2	48 (4)
3-4	23 (2)
Missing data	497 (40)
BMI, n (%)	
<18.5	32 (3)
18.5-25	446 (36)
25-30	409 (33)
>30	266 (22)
Missing data	76 (6)
Comorbidities, n (%)	
Creatinine clearance <40 mL/minute	57 (5)
Heart failure	105 (9)
Coronary artery disease	123 (10)
Chronic obstructive pulmonary disease	36 (3)
Diabetes	174 (14)

Table 2. Cancer characteristics of the initial cohort (N=1229).

Tumor size					
Status	T1	T2	T3	T4	Missing data
Participants, n (%)	567 (46)	286 (23)	36 (3)	250 (20)	90 (7)
Lymph nodes					

Tumor size					
Status	N0	N1	N2	N3	Missing data
Participants, n (%)	614 (50)	243 (20)	55 (4)	55 (4)	262 (21)
Grade SBR ^a					
Status	I	II		III	Missing data
Participants, n (%)	188 (15)	648 (53)		281 (23)	112 (9)
Estrogen receptor					
Status	Positive		Negative		Missing data
Participants, n (%)	978 (80)		145 (12)		106 (9)
Progesterone receptor					
Status	Positive		Negative		Missing data
Participants, n (%)	838 (68)		285 (23)		106 (9)

^aSBR: Scarff-Bloom-Richardson.

Patient demographics and characteristics were evaluated on the date of breast cancer diagnosis (Table 1). Median age was 75 years (range: 70-100 years), with 317/1229 (26%) patients aged 80 years or older. Performance status was generally good, as most are categorized as 0 or 1. The main comorbidities were diabetes (174/1229 patients, or 14%), followed by coronary artery disease (123/1229 patients, 10%) and cardiac insufficiency (105/1229 patients, 9%).

The majority presented early-stage tumors (T1 in 567/1229 patients, with a prevalence of 46%), and lymph node involvement was mostly absent (N0 in 614/1229 patients, or 50%). The tumors were typically SBR grade II and 80% (978/1229 patients) were ER-positive. Progesterone receptor positivity was also high at 68% (838/1229 patients). Twelve percent of patients (149/1229) were reported to have received chemotherapy (Table 2).

Development Cohort

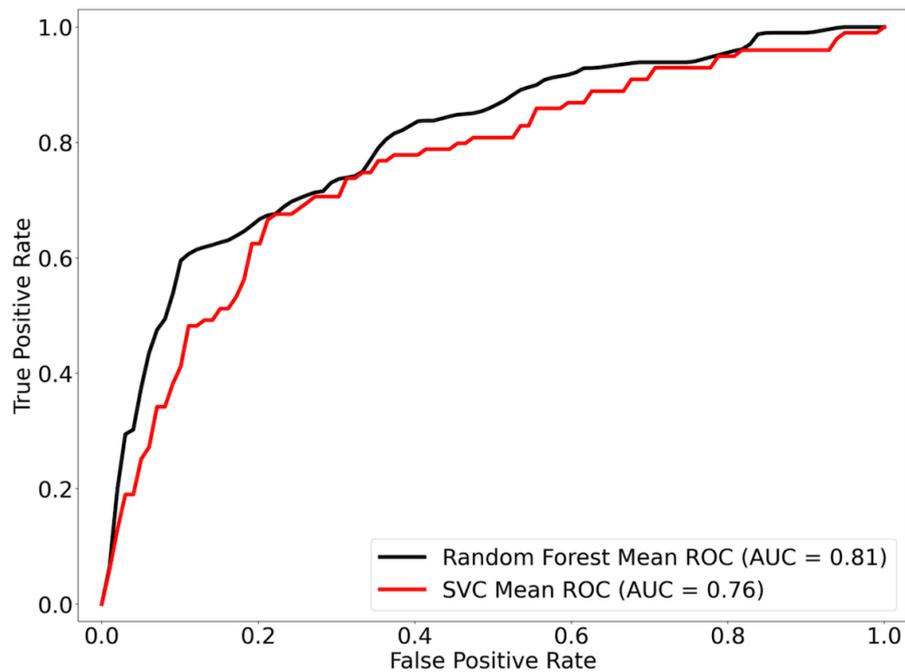
The final cohort was divided into “reference” and “model” cohorts for model development (Figure 2). A total of

327 patients, that is, 50% of patients meeting the criteria for manifold-estimated derivation training were randomly selected. The purpose of this training group was to calibrate the manifold-estimated derivation–scoring algorithm. The model data comprised all remaining patients (466, 59% of the model development cohort).

Features Performance and Area Under the Curve Scores

In Figure 3, we ascertained the predictive efficacy of the selected variables using RFC and SVC. Analyzing the receiver operating characteristic curves, both models demonstrated commendable predictive capabilities. RFC yielded a mean area under the curve (AUC) of 0.81 (SD 0.06) and a mean accuracy of 0.82 (SD 0.02), while SVC followed closely with a mean AUC of 0.76 (SD 0.05) and a mean accuracy of 0.78 (SD 0.01). The overlapping SDs of these scores suggest that the differences in their performance are not statistically significant.

Figure 3. Receiver operating characteristic curves for 5-year mortality predictive models. The predictive efficacy of the selected features was ascertained using Random Forest Classification and Support Vector Classifier. Results are presented as the mean of ROC and AUC values derived from 5-fold cross-validation. AUC: area under the curve; ROC: receiver operating characteristic; SVC: Support Vector Classifier.



The overall relative importance of variables for the prediction of the 5-year outcome was also determined by RFC (Table 3). Age, tumor size, and hemoglobin were the top predictors, closely followed by lymphocyte count and BMI. Curiously, the cancer grade, axillary lymph nodes involvement, and the presence of comorbidities ranked low in overall importance.

This indicates that although typically taken as important factors from a clinical perspective, comorbidities and cancer grade alone are not the best prognostic features in a patient; rather, a patient's overall biological profile may be more valuable, underscoring the usefulness of manifold learning as a prognostic tool.

Table 3. Overall importance of predictors according to Random Forest Classification.

Variable	Importance (%)
Age	18.33
Tumor size	17.26
Hemoglobin (g/dL)	16.41
Lymphocytes (g/L)	14.84
BMI	13.06
Lymph nodes involvement	10.39
SBR ^a grade	4.06
ER ^b status	2.88
Comorbidities	2.78

^aSBR: Scarff-Bloom-Richardson.

^bER: estrogen receptor.

Model Stability

Patients in the test and model groups from the same cluster of origin were compared to evaluate whether data points would exhibit similar distributions (appear in proximity to each other) in the new manifold space across 10 different manifold initializations. The distributions of the test group (n=140) consistently matched closely with those of the model group (n=326), with all *P* values being above the threshold of .05 indicating a lack of significant variation between groups (Figure S1 in Multimedia Appendix 1).

Prognostic Ability

The primary objective of our study is to evaluate the prognostic ability of the manifold learning model, as measured by the 5-year survival rate of our population. The 3D clusters in Figure 1 illuminated the landscape of our dataset, representing local groups of patients characterized by distinct clinical and prognostic profiles. Clusters are colored based on the overall mortality rate of included patients: Groups 0, 1, and 4 in green have the best prognosis with a

5-year survival rate of more than 80% while group 3 has the worst prognosis with a 5-year mortality rate of at least 35% (in mm), and median age, underscoring the diversity in our cohort.

Table 4 further elucidates the variability in values across the patient clusters, especially in BMI, tumor size

Table 4. Characteristics of the 6 clusters defined by manifold learning.

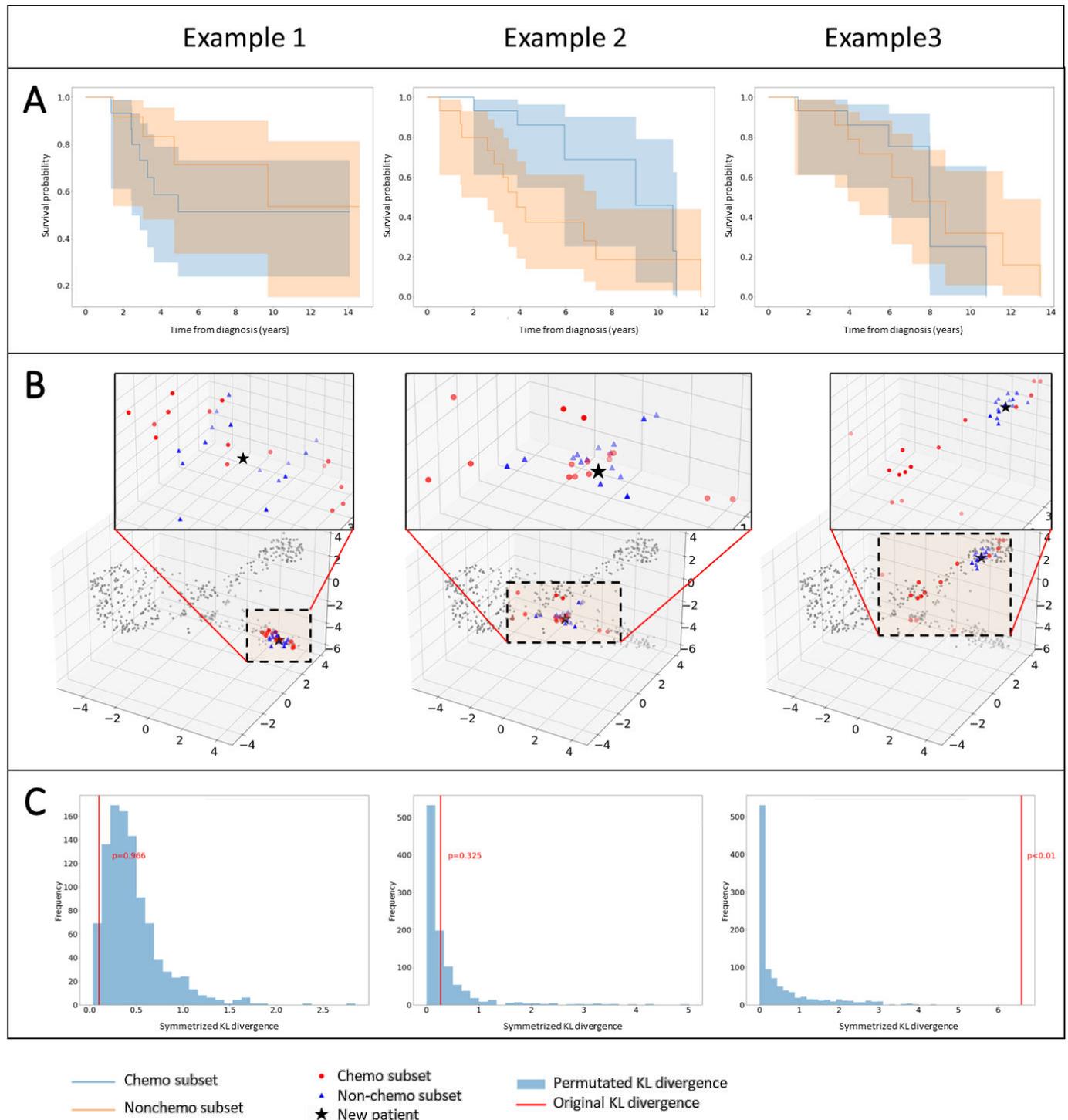
Feature	Cluster					
	0	1	2	3	4	5
Hemoglobin (g/dL)	13.4	13.3	13	13	13.3	11.9
BMI	25	28.4	24.9	28.9	25.6	23
Lymph nodes involved	0.6	0.8	1.6	6.2	1	0.8
Tumor size (mm)	19.6	19.3	26.1	65.8	23.1	30.6
Age (years)	75.8	76.3	77.8	77.4	79.2	80.5
Lymphocytes (g/L)	1.8	2.1	2.1	1.6	3.4	1.6
Comorbidities	0	1	0.4	0.3	0	0.9
Estrogen receptor status	1	1	0	0.8	1	1
SBR ^a (high/low)	0	0.1	0.7	0.5	0.8	0.6

^aSBR: Scarff-Bloom-Richardson.

Predictive Ability

Next, we attempted to ascertain the individual benefit of performing adjuvant chemotherapy, demonstrated in [Figure 4](#) with 3 examples.

Figure 4. Three case examples assessing the individual benefit of adjuvant chemotherapy. (A) The closest chemotherapy-treated and non-chemotherapy-treated patients to a new patient are identified in the 3D manifold and their survival curves are compared to show the treatment’s potential benefit or lack thereof. (B) The new patient’s position in the 3D manifold (black star), with the 15 closest patients of each treatment groups are shown, displaying varying distributions of treatment subgroups. (C) To quantify distances between the subgroups, the real calculated KL divergence between the treatment groups’ distributions (red line) was compared with that of permuted data (blue histograms) to verify whether observed divergences between treatment subgroups are significant or not. KL: Kullback-Leibler.



When a target patient is localized in the 3D manifold, the closest patient profiles are identified. This is done for 2 treatment groups based on whether the patients received chemotherapy (chemo and nonchemo groups), permitting the visualization of KM survival curves that would show the treatment’s potential benefit or lack thereof (Figure 4A).

Figure 4B shows the target patient’s position in the 3D manifold (black star), with the 15 closest patients of each treatment groups also marked. In examples 1 and 2, the 2 treatment groups are found to be well “mixed” in the local vicinity of the target, indicating that the target profile is well represented by similar chemotherapy-treated and non-chemotherapy-treated patients. To quantify distances between

the subgroups, we used permutation analysis (Figure 4C). The real calculated KL divergences between the treatment groups' distributions (red line) for examples 1 and 2 fall well within the range of what could be expected by chance (blue histograms) ($P>.1$), indicating that the observed divergences are not significant.

Example 3 showcases a situation where patients from the 2 treatment groups are not well mixed in the local vicinity of the target patient. In this case, the real KL divergence is far right of the histogram ($P<.01$), suggesting a significant difference between the distributions. Thus, the KM survival curves and any conclusion drawn from them must be taken with consideration of the heterogeneity in the profiles of the treatment groups being compared.

Discussion

Principal Findings

From an initial cohort of 1229 patients, we used 793 (65%) to develop a model that clustered patients by their clinical and biological features. These clusters represent a potential prognostic tool for physicians, attributing a risk of mortality in 5 years to patients with consideration to multivariate profiles. The model is further able to indicate the potential benefit or lack thereof of chemotherapy treatment in older adult patients. We found that the predictors used in our model gave a good overall result of 0.81 and 0.76 AUCs with RFC and SVC, respectively.

In summation, our multifaceted approach, blending manifold learning with classical machine learning paradigms and intuitive data visualizations, has unveiled profound insights into the prognosis determinants of early-stage breast cancer in older adults. These revelations bring a more nuanced understanding of the disease and hold promise for tailoring patient-specific therapeutic strategies. Our study's utilization of manifold learning and advanced machine learning algorithms represents a significant contribution to oncology. The accuracy of 81% in differentiating patient subgroups through manifold learning is impressive, showcasing an advancement beyond traditional linear models [33]. This approach is in line with recent trends in personalized medicine [34,35], which discuss the potential of machine learning in cancer prognosis. The high AUC values achieved by RFC and SVC reflect the importance of our combined predictors in medical diagnostics, aligning with the findings of recent studies on the application of machine learning in cancer detection [36,37]. The application of data visualization techniques such as heatmaps and 3D scatterplots in elucidating complex clinical relationships is noteworthy. This approach is supported by advancements in data visualization in medical research, as seen in the study by Borkin et al [38] on how data visualization supports medical decision-making [39,40].

Limitations

The present results should be interpreted in the light of some limitations. First, the monocentric nature of the research

may impact the representativeness of the cohort, potentially affecting the generalizability of our findings. Second, the exclusion of specific patient characteristics, such as the ONCODAGE score [41], from our datasets may have limited the comprehensiveness of our prognostic tools. Third, the retrospective design of the study constrains our ability to establish causality between clinical characteristics and patient outcomes. A fourth limitation concerns the fact that patients may present with or have a history of multiple comorbidities. We chose to group together patients with any number of comorbidities for reasons related to (1) the reduction of the sample size for each category of comorbidity, and (2) the potential skewing of patient distribution in the 3D manifold due to multiple related qualitative variables. PaCMAP is susceptible to "overseparate" the population if provided with too many binary features. These reasons in mind, we nonetheless acknowledge that omitting the consideration of multiple comorbidities is a limitation of the study. Other notable limitations include the absence of cancer-specific or treatment-specific survival metrics, a lack of detailed analysis on specific comorbidities, and the need for more data to enhance the less populated clusters. Furthermore, the external validation of our model remains pending, which is crucial for assessing its generalizability.

Future Prospects

Looking forward, the promising application of manifold learning in oncology, as demonstrated in our study, aligns with the burgeoning field of personalized medicine. The integration of machine learning in personalized cancer therapy, as discussed by Danishuddin et al [42], supports the potential of such approaches. The development of advanced AI-driven prognostic tools, particularly for older adult patients who are often underrepresented in clinical trials, could revolutionize treatment guidelines and care approaches. The rapid advancement of machine learning techniques poses a challenge in ensuring the longevity and relevance of models, necessitating continuous updates. This is echoed in the broader context of AI in health care, as discussed in Topol's [43] comprehensive review of AI in medicine. Concerns about the adoption of AI tools due to accuracy, explainability, and ethical considerations are also prevalent, as reflected in the exploration of implementing AI in clinical practice by Char et al [44]. Our findings may open up avenues for the personalized treatment specifically catered to neglected populations in oncology, starting with geriatric patients with breast cancer. We expect our software to provide rapid guidance to physicians in the process of charting treatment plans for their patients, going beyond simple monovariate statistics and instead considering patients' combined clinical and biological profiles.

Conclusions

Our study aimed to further the management of early breast cancer in older adult patients by integrating cutting-edge AI techniques. We proposed a technique that uses patient data to create a visualizable 3D map of pathology profiles that allow rapid prognostic estimations for new patients. These prognostic predictions include the potential

benefits of treatment strategies such as chemotherapy, aiding clinical decision-making. It reflects the ongoing evolution in oncology, emphasizing the importance of tailored treatment strategies and highlighting both the potential and the challenges of AI applications in health care. This study also prompts considerations for future research directions and ethical implications in the rapidly evolving field of AI in medicine.

Data Availability

The data that support the findings of this study are not publicly available due to privacy and confidentiality agreements. However, pseudonymized data may be made available upon reasonable request to the corresponding author and following approval by the French Commission Nationale de l'Informatique et des Libertés under the data-sharing agreements of the Léon Bérard Cancer Center.

Authors' Contributions

FR and PH contributed to conceptualization, methodology, project administration, validation, writing (review and editing). MA contributed to methodology, investigation, data curation, validation, and writing (original draft and editing). AA contributed to conceptualization, supervision, writing (review and editing), and resources.

Conflicts of Interest

The model described in this study was developed by GeodAISics, who are listed among the authors. PH reports : grants; personal fees; and nonfinancial support from PFIZER, LILLY, DAICHII, and ASTRAZENECA grants, and nonfinancial support from NOVARTIS and ROCHE—personal fees and nonfinancial support from SEAGEN, GILEAD, and MSD—cofounder and chief medical officer of GEODASICS. AA is the founder and CEO of GeodAISics. The remaining authors have no conflicts of interest.

Multimedia Appendix 1

Stability analysis of manifold learning applied to clustered data. The original cohort data were divided into 2 groups; 70% of every cluster was pooled into the model group, and the remaining 30% was pooled into the test group. A fresh manifold learning process was applied to the model group, and the test group was then projected onto the newly generated manifold. Patients in the test and model groups from the same cluster of origin were compared to evaluate whether they would exhibit similar distributions (appear in proximity to each other) in the new manifold space. (A) Examples of permutation analysis of clusters 0 and 1. The permutation test determined whether the observed KL (red line) divergence was significantly different from what can be expected from random shuffling of the 2 groups (blue histograms). (B) Table summarizing the median P values of the stability tests across 10 different manifold initializations. All P values above .05 indicated a lack of significant variation between groups.

[[PNG File \(Portable Network Graphics File\), 214 KB-Multimedia Appendix 1](#)]

References

1. Howlader N, Noone AM, Krapcho M, et al. SEER cancer statistics review, 1975-2017. Based on November 2019 SEER data submission, posted to the SEER web site. Cronin KA, editor. National Cancer Institute; URL: https://seer.cancer.gov/archive/csr/1975_2017/ [Accessed 2025-05-04]
2. American Cancer Society. Breast cancer facts & figures 2019-2020. URL: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2019-2020.pdf> [Accessed 2025-05-04]
3. Ferlay J, Ervik M, Lam F, et al. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer; 2020. URL: <https://gco.iarc.fr/today> [Accessed 2025-05-20]
4. Tesarova P. Specific aspects of breast cancer therapy of elderly women. *Biomed Res Int*. 2016;2016:1381695. [doi: [10.1155/2016/1381695](https://doi.org/10.1155/2016/1381695)] [Medline: [27807536](https://pubmed.ncbi.nlm.nih.gov/27807536/)]
5. Glaser R, Marinopoulos S, Dimitrakakis C. Breast cancer treatment in women over the age of 80: a tailored approach. *Maturitas*. Apr 2018;110:29-32. [doi: [10.1016/j.maturitas.2018.01.014](https://doi.org/10.1016/j.maturitas.2018.01.014)] [Medline: [29563032](https://pubmed.ncbi.nlm.nih.gov/29563032/)]
6. Clarke M, Coates AS, et al. Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. *Lancet*. Jan 5, 2008;371(9606):29-40. [doi: [10.1016/S0140-6736\(08\)60069-0](https://doi.org/10.1016/S0140-6736(08)60069-0)] [Medline: [18177773](https://pubmed.ncbi.nlm.nih.gov/18177773/)]
7. Peto R, Davies C, Early Breast Cancer Trialists' Collaborative Group (EBCTCG), et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet*. Feb 4, 2012;379(9814):432-444. [doi: [10.1016/S0140-6736\(11\)61625-5](https://doi.org/10.1016/S0140-6736(11)61625-5)] [Medline: [22152853](https://pubmed.ncbi.nlm.nih.gov/22152853/)]
8. Buchanan BG, Shortlife EH, editors. An expert system for oncology protocol management. In: Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading: Addison-Wesley; 1984:876-881.
9. Lieber J, Bresson B. Case-based reasoning for breast cancer treatment decision helping. In: Blanzieri E, Portinale L, editors. Advances in Case-Based Reasoning EWCBR 2000 Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence). Vol 1898. Springer; 2000:173-185.

10. Hawley ST, Li Y, An LC, et al. Improving breast cancer surgical treatment decision making: the iCanDecide randomized clinical trial. *J Clin Oncol*. Mar 1, 2018;36(7):659-666. [doi: [10.1200/JCO.2017.74.8442](https://doi.org/10.1200/JCO.2017.74.8442)] [Medline: [29364772](https://pubmed.ncbi.nlm.nih.gov/29364772/)]
11. Wyatt KD, Jenkins SM, Plevak MF, Venegas Pont MR, Pruthi S. A personalized, web-based breast cancer decision making application: a pre-post survey. *BMC Med Inform Decis Mak*. Oct 21, 2019;19(1):196. [doi: [10.1186/s12911-019-0924-7](https://doi.org/10.1186/s12911-019-0924-7)] [Medline: [31638964](https://pubmed.ncbi.nlm.nih.gov/31638964/)]
12. Xu F, Sepúlveda MJ, Jiang Z, et al. Artificial intelligence treatment decision support for complex breast cancer among oncologists with varying expertise. *JCO Clin Cancer Inform*. Aug 2019;3:1-15. [doi: [10.1200/CCI.18.00159](https://doi.org/10.1200/CCI.18.00159)] [Medline: [31419181](https://pubmed.ncbi.nlm.nih.gov/31419181/)]
13. Mazo C, Kearns C, Mooney C, Gallagher WM. Clinical decision support systems in breast cancer: a systematic review. *Cancers (Basel)*. Feb 6, 2020;12(2):369. [doi: [10.3390/cancers12020369](https://doi.org/10.3390/cancers12020369)] [Medline: [32041094](https://pubmed.ncbi.nlm.nih.gov/32041094/)]
14. Wishart GC, Azzato EM, Greenberg DC, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res*. 2010;12(1):R1. [doi: [10.1186/bcr2464](https://doi.org/10.1186/bcr2464)] [Medline: [20053270](https://pubmed.ncbi.nlm.nih.gov/20053270/)]
15. Jiang X, Wells A, Brufsky A, Neapolitan R. A clinical decision support system learned from data to personalize treatment recommendations towards preventing breast cancer metastasis. *PLoS ONE*. 2019;14(3):e0213292. [doi: [10.1371/journal.pone.0213292](https://doi.org/10.1371/journal.pone.0213292)] [Medline: [30849111](https://pubmed.ncbi.nlm.nih.gov/30849111/)]
16. Hughes KS, Schnaper LA, Berry D, et al. Lumpectomy plus tamoxifen with or without irradiation in women 70 years of age or older with early breast cancer. *N Engl J Med*. Sep 2, 2004;351(10):971-977. [doi: [10.1056/NEJMoa040587](https://doi.org/10.1056/NEJMoa040587)] [Medline: [15342805](https://pubmed.ncbi.nlm.nih.gov/15342805/)]
17. Clough-Gorr KM, Thwin SS, Stuck AE, Silliman RA. Examining five- and ten-year survival in older women with breast cancer using cancer-specific geriatric assessment. *Eur J Cancer*. Apr 2012;48(6):805-812. [doi: [10.1016/j.ejca.2011.06.016](https://doi.org/10.1016/j.ejca.2011.06.016)] [Medline: [21741826](https://pubmed.ncbi.nlm.nih.gov/21741826/)]
18. Stotter A, Reed MW, Gray LJ, Moore N, Robinson TG. Comprehensive Geriatric Assessment and predicted 3-year survival in treatment planning for frail patients with early breast cancer. *Br J Surg*. Apr 2015;102(5):525-533; [doi: [10.1002/bjs.9755](https://doi.org/10.1002/bjs.9755)] [Medline: [25708660](https://pubmed.ncbi.nlm.nih.gov/25708660/)]
19. Alaa AM, Gurdasani D, Harris AL, Rashbass J, van der Schaar M. Machine learning to guide the use of adjuvant therapies for breast cancer. *Nat Mach Intell*. 2012;3(8):716-726. [doi: [10.1038/s42256-021-00353-8](https://doi.org/10.1038/s42256-021-00353-8)]
20. Shachar SS, Muss HB. Internet tools to enhance breast cancer care. *NPJ Breast Cancer*. 2016;2:16011. [doi: [10.1038/npjbcancer.2016.11](https://doi.org/10.1038/npjbcancer.2016.11)] [Medline: [28721377](https://pubmed.ncbi.nlm.nih.gov/28721377/)]
21. van Maaren MC, van Steenbeek CD, Pharoah PDP, et al. Validation of the online prediction tool PREDICT v. 2.0 in the Dutch breast cancer population. *Eur J Cancer*. Nov 2017;86:364-372. [doi: [10.1016/j.ejca.2017.09.031](https://doi.org/10.1016/j.ejca.2017.09.031)] [Medline: [29100191](https://pubmed.ncbi.nlm.nih.gov/29100191/)]
22. Van der Schaar Lab. Adjutorium: a machine learning framework for personalized survival prediction. URL: <https://vanderschaar-lab.com/adjutorium/> [Accessed 2025-05-06]
23. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol*. Feb 15, 2001;19(4):980-991. [doi: [10.1200/JCO.2001.19.4.980](https://doi.org/10.1200/JCO.2001.19.4.980)] [Medline: [11181660](https://pubmed.ncbi.nlm.nih.gov/11181660/)]
24. de Glas NA, van de Water W, Engelhardt EG, et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. *Lancet Oncol*. Jun 2014;15(7):722-729. [doi: [10.1016/S1470-2045\(14\)70200-1](https://doi.org/10.1016/S1470-2045(14)70200-1)] [Medline: [24836274](https://pubmed.ncbi.nlm.nih.gov/24836274/)]
25. Engelhardt EG, Garvelink MM, de Haes JHCJM, et al. Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models. *J Clin Oncol*. Jan 20, 2014;32(3):238-250. [doi: [10.1200/JCO.2013.50.3417](https://doi.org/10.1200/JCO.2013.50.3417)] [Medline: [24344212](https://pubmed.ncbi.nlm.nih.gov/24344212/)]
26. van der Plas-Krijgsman WG, Giardiello D, Putter H, et al. Development and validation of the PORTRET tool to predict recurrence, overall survival, and other-cause mortality in older patients with breast cancer in the Netherlands: a population-based study. *Lancet Healthy Longev*. Nov 2021;2(11):e704-e711. [doi: [10.1016/S2666-7568\(21\)00229-4](https://doi.org/10.1016/S2666-7568(21)00229-4)] [Medline: [36098027](https://pubmed.ncbi.nlm.nih.gov/36098027/)]
27. Heudel P, Livartowski A, Arveux P, Willm E, Jamain C. The ConSoRe project supports the implementation of big data in oncology. *Bull Cancer*. Nov 2016;103(11):949-950. [doi: [10.1016/j.bulcan.2016.10.001](https://doi.org/10.1016/j.bulcan.2016.10.001)] [Medline: [27816168](https://pubmed.ncbi.nlm.nih.gov/27816168/)]
28. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. *CA Cancer J Clin*. Mar 2017;67(2):93-99. [doi: [10.3322/caac.21388](https://doi.org/10.3322/caac.21388)] [Medline: [28094848](https://pubmed.ncbi.nlm.nih.gov/28094848/)]
29. Wolff AC, Hammond MEH, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol*. Nov 1, 2013;31(31):3997-4013. [doi: [10.1200/JCO.2013.50.9984](https://doi.org/10.1200/JCO.2013.50.9984)] [Medline: [24101045](https://pubmed.ncbi.nlm.nih.gov/24101045/)]
30. Wang Y. GitHub repository. PaCMAP: Pairwise Controlled Manifold Approximation. URL: <https://github.com/YingfanWang/PaCMAP> [Accessed 2025-05-06]

31. Scikit-learn developers. "sklearn.cluster.MeanShift," scikit-learn documentation. URL: <https://scikit-learn.org/dev/modules/generated/sklearn.cluster.MeanShift.html> [Accessed 2025-05-06]
32. Davidson-Pilon C. Lifelines: survival analysis in Python," Read the Docs. URL: https://antiphishing.vadesecond.com/v4?f=WExObUdsNkxHUTVwMTdKdg9XvgXO8Cu4fwIkJ7gAD8xPip7tYbm7MrM9TOzJlbbv&i=dktEMmMyTnlid1lsUjVvYvDN_JOcSy-MnYrzg1_2L5o&k=dPpv&r=bGMwQ1dycHZ0bUpyOWJIRhJnIftWtEMd5uOOEiOFp19wG5_6SGkO_QGtssB6mbs-Rakx_9mw4s4NHld5ghzXZQ&s=ea1df091c2a129d37187a974face80fa62a90968c96c3a752e4cddb3d1ee5183&u=https%3A%2F%2Flifelines.readthedocs.io%2Fen%2Flatest%2F [Accessed 2025-05-06]
33. Bhavnani SK, Zhang W, Visweswaran S, Raji M, Kuo YF. A framework for modeling and interpreting patient subgroups applied to hospital readmission: visual analytical approach. *JMIR Med Inform.* Dec 7, 2022;10(12):e37239. [doi: [10.2196/37239](https://doi.org/10.2196/37239)] [Medline: [35537203](https://pubmed.ncbi.nlm.nih.gov/35537203/)]
34. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J.* 2015;13(13):8-17. [doi: [10.1016/j.csbj.2014.11.005](https://doi.org/10.1016/j.csbj.2014.11.005)] [Medline: [25750696](https://pubmed.ncbi.nlm.nih.gov/25750696/)]
35. Ozer ME, Sarica PO, Arga KY. New machine learning applications to accelerate personalized medicine in breast cancer: rise of the support vector machines. *OMICS.* May 2020;24(5):241-246. [doi: [10.1089/omi.2020.0001](https://doi.org/10.1089/omi.2020.0001)] [Medline: [32228365](https://pubmed.ncbi.nlm.nih.gov/32228365/)]
36. Shafiq A, Çolak AB, Sindhu TN, Lone SA, Abushal TA. Modeling and survival exploration of breast carcinoma: A statistical, maximum likelihood estimation, and artificial neural network perspective. *Artificial Intelligence in the Life Sciences.* Dec 2023;4:100082. [doi: [10.1016/j.ailsci.2023.100082](https://doi.org/10.1016/j.ailsci.2023.100082)]
37. Wu R, Luo J, Wan H, et al. Evaluation of machine learning algorithms for the prognosis of breast cancer from the Surveillance, Epidemiology, and End Results database. *PLoS ONE.* 2023;18(1):e0280340. [doi: [10.1371/journal.pone.0280340](https://doi.org/10.1371/journal.pone.0280340)] [Medline: [36701415](https://pubmed.ncbi.nlm.nih.gov/36701415/)]
38. Borkin MA, Vo AA, Bylinskii Z, et al. What makes a visualization memorable? *IEEE Trans Vis Comput Graph.* Dec 2013;19(12):2306-2315. [doi: [10.1109/TVCG.2013.234](https://doi.org/10.1109/TVCG.2013.234)] [Medline: [24051797](https://pubmed.ncbi.nlm.nih.gov/24051797/)]
39. Palmas G, Bachynskiy M, Oulasvirta A, Seidel HP, Weinkauff T. MovExp: a versatile visualization tool for human-computer interaction studies with 3D performance and biomechanical data. *IEEE Trans Vis Comput Graph.* Dec 2014;20(12):2359-2368. [doi: [10.1109/TVCG.2014.2346311](https://doi.org/10.1109/TVCG.2014.2346311)] [Medline: [26356950](https://pubmed.ncbi.nlm.nih.gov/26356950/)]
40. Park S, Bekemeier B, Flaxman A, Schultz M. Impact of data visualization on decision-making and its implications for public health practice: a systematic literature review. *Inform Health Soc Care.* Apr 3, 2022;47(2):175-193. [doi: [10.1080/17538157.2021.1982949](https://doi.org/10.1080/17538157.2021.1982949)] [Medline: [34582297](https://pubmed.ncbi.nlm.nih.gov/34582297/)]
41. Soubeyran P, Bellera C, Goyard J, et al. Screening for vulnerability in older cancer patients: the ONCODAGE Prospective Multicenter Cohort Study. *PLoS One.* 2014;9(12):e115060. [doi: [10.1371/journal.pone.0115060](https://doi.org/10.1371/journal.pone.0115060)] [Medline: [25503576](https://pubmed.ncbi.nlm.nih.gov/25503576/)]
42. Danishuddin A, Khan S, Kim JJ. From cancer big data to treatment: artificial intelligence in cancer research. *J Gene Med.* Jan 2024;26(1):e3629. [doi: [10.1002/jgm.3629](https://doi.org/10.1002/jgm.3629)] [Medline: [37940369](https://pubmed.ncbi.nlm.nih.gov/37940369/)]
43. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* Jan 2019;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [Medline: [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/)]
44. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med.* Mar 15, 2018;378(11):981-983. [doi: [10.1056/NEJMp1714229](https://doi.org/10.1056/NEJMp1714229)] [Medline: [29539284](https://pubmed.ncbi.nlm.nih.gov/29539284/)]

Abbreviations

- AI:** artificial intelligence
- AUC:** area under the curve
- COPD:** chronic obstructive pulmonary disease
- ER:** estrogen receptor
- HER2:** human epidermal growth factor receptor 2
- KL:** Kullback-Leibler
- KM:** Kaplan-Meier
- PaCMAP:** Pairwise Controlled Manifold Approximation
- PR:** progesterone receptor
- RFC:** Random Forest Classification
- SBR:** Scarff-Bloom-Richardson
- SVC:** Support Vector Classification

Edited by Naomi Cahill; peer-reviewed by James C L Chow, Zhu Ping; submitted 05.07.2024; final revised version received 22.01.2025; accepted 24.02.2025; published 23.05.2025

Please cite as:

Heudel P, Ahmed M, Renard F, Attye A

Leveraging Digital Twins for Stratification of Patients with Breast Cancer and Treatment Optimization in Geriatric Oncology: Multivariate Clustering Analysis

JMIR Cancer 2025;11:e64000

URL: <https://cancer.jmir.org/2025/1/e64000>

doi: [10.2196/64000](https://doi.org/10.2196/64000)

© Pierre Heudel, Mashal Ahmed, Felix Renard, Arnaud Attye. Originally published in JMIR Cancer (<https://cancer.jmir.org>), 23.05.2025. 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.