ecancermedicalscience

Predictors of early relapse among breast cancer patients in Sudan: a Cox Proportional Hazards Model approach

Mayada Mustafa Ahmed[1a], Bashier Ibrahim Osman[1b], Alnada Abdalla Mohamed Ibrahim[2c], Maisa Elfadul[3d], Mounkaila Noma[4e], Kamal Eldein H Mohamed[5f] and Nazik Elmalaika Obaid Seid Ahmed Husain[6g]

[1]Department of Pharmacology, Faculty of Pharmacy, University of Khartoum, 1111 Al Qasr Avenue, Khartoum B1996, Sudan
[2]Department of Pharmaceutics, Faculty of Pharmacy, University of Khartoum, 1111 Al Qasr Avenue, Khartoum B1996, Sudan
[3]Public and Tropical Health Programs, Epidemiology and Biostatistics, University of Medical Sciences and Technology, Sudan Medical Specialisation Board, 11111 Hospital Street, P 102 Khartoum, Sudan
[4]Medical and Health Sciences/Biostatistics, University of Medical Sciences & Technology, 1111 Khartoum B12810, Sudan
[5]Khartoum Oncology Hospital, 12345 El Qasr Avenue, Khartoum B 846, Sudan
[6]Department of Pathology, Faculty of Medicine, University of Tabuk, Tabuk 71491, Saudi Arabia
[a] https://orcid.org/0000-0001-5632-8628
[b] https://orcid.org/0000-0002-4931-7098
[c] https://orcid.org/0000-0002-4604-4653
[d] https://orcid.org/0000-0003-2976-2642
[e] https://orcid.org/0000-0003-1772-2686
[f] https://orcid.org/0000-0003-0800-0454
[g] https://orcid.org/0000-0001-8333-5735

Research

Abstract

Background: Integrating molecular markers with clinical parameters into prognostic models enhances the prediction of therapy response, recurrence, and mortality. However, improved prognostication in breast cancer has not been well established in Sudan. This study aims to construct and validate a data-driven prognostic index (PI) to inform chemotherapy decisions for breast cancer patients in Sudan.

Methods: A prospective cohort of 257 breast cancer patients treated with neoadjuvant chemotherapy at Khartoum Oncology Hospital was followed to identify factors associated with early relapse (≤18 months). Variables included clinical variables (age, tumour stage, lymph node (LN) status and neoadjuvant tumour response), as well as immunohistochemical markers (Human Epidermal Growth Factor Receptor 2, Ki-67 and Topoisomerase 2A). The PI was constructed using Cox Proportional Hazards regression. Model performance was assessed using Kaplan–Meier survival analysis, receiver operating characteristic curve and calibration test. The clinical utility of the model was evaluated using the net reclassification improvement (NRI), and performance was compared with the Nottingham Prognostic Index (NPI).

Results: Initial LN involvement was the only factor statistically associated with early relapse (p = 0.004; HR = 3.79, 95% CI: 1.53–9.37). A seven-factor PI was developed, stratifying patients into five distinct risk groups. Kaplan–Meier survival analysis revealed a significant difference between the groups (log-rank p < 0.0001). The PI demonstrated

Correspondence to: Mayada Mustafa Ahmed
Email: Mayada99mustafa@gmail.com

*e*cancer 2026, **20**:2161
https://doi.org/10.3332/ecancer.2026.2161

Published: 08/07/2026
Received: 06/01/2026

Publication costs for this article were supported by *e*cancer (UK Charity number 1176307).

good calibration (p > 0.05) and higher predictive accuracy than the NPI, with an area under the curve of 0.81 (95% CI: 0.717–0.895) compared to 0.74 (95% CI: 0.643–0.834) for the NPI. NRI was 0.13.

Conclusion: The developed PI demonstrated strong predictive performance and robust internal validation for predicting early relapse in Sudanese breast cancer patients. It offers a promising framework for personalised treatment decisions in resource-limited healthcare settings.

Keywords: *breast cancer, early relapse, prognostic index, Cox regression, lymph node, Sudan*

Background

Breast cancer is the most common cancer among women worldwide and a leading cause of cancer-related mortality [1, 2]. Its clinical presentation is highly heterogeneous, with diverse biological and molecular characteristics that result in variable patients, outcomes [3, 4]. While advances in screening and treatment have significantly improved survival rates in developed countries, outcomes in many developing nations, including Sudan, remain suboptimal [4–8].

Recurrence and metastasis are particularly challenging in high-risk populations, and despite significant advances in understanding the disease, breast cancer management remains complex, especially at advanced stages [5, 9, 10]. Prognosis – the likelihood of disease recurrence or death – is crucial in guiding treatment decisions [3, 11]. The Nottingham Prognostic Index (NPI) remains a widely used validated prognostic tool in breast cancer settings that offers robust long-term prognostication. However, its use is increasingly being complemented by molecular-based prognostic models. Despite their growing importance, many of these models lack validation in specific populations, particularly young patients age and aggressive disease subtypes [12–14]. Therefore, concerns have been raised regarding the use of some of the recently introduced Multigene prognostic genomic assays, especially in premenopausal and BRCA1/2 breast cancer patients [14].

Consequently, many studies validated these new tools in other populations distinct from those of European ancestry. A performance gap was observed in many studies involving certain Asian populations [13–17]. Furthermore, although the performance and clinical utility of these tools in sub-Saharan Africa have not been elucidated, the same gap is expected to exist because these tools do not fully capture the diverse prognostic factors in breast cancer in these populations. Thus, several measures had to be addressed to mitigate the exacerbation of this performance gap by developing and validating more reliable population - specific tools [13, 16, 17].

In Sudan, chemotherapy remains the cornerstone of breast cancer treatment. Treatment decisions often rely on generalised models not validated for local conditions. Limited resources and a lack of individualised treatment protocols contribute to a high rate of early relapse [2, 4, 8]. Despite the global progress in understanding breast cancer prognosis, there is a lack of a validated population – specific tool in Sudan, particularly in predicting early relapse. Moreover, advanced genomic tools such as Oncotype DX, MammaPrint and Pam50 are inaccessible, unaffordable and require further validation. This highlights the need for a cost-effective, locally derived prognostic tool.

This study aims to establish and validate a data-driven prognostic index (PI) and compare its predictive performance with the NPI, to improve therapeutic guidance within a resource-limited oncology setting.

Methods

Study design

The study followed TRIPOD guidelines for prognostic model development and validation, ensuring transparent reporting and reproducibility [18].

A prospective cohort study design was employed to ensure high-quality longitudinal data on early relapse in breast cancer patients. This hospital-based study was conducted at Khartoum Oncology Hospital (KOH) in Khartoum, Sudan, under the oversight of the institutional review board (IRB) subcommittee. To ensure methodological rigour, the committee utilised reference frameworks and provided multidisciplinary guidance from oncologists, pharmacologists and biostatisticians. [18–24].

Study population

Patients with stage I to III breast cancer, who were receiving neoadjuvant chemotherapy courses in the period from 2018 to 2019, were assessed for their eligibility to participate in the study. Inclusion criteria comprised patients aged 18 years or older with complete histopathological data. Patients with inadequate or poor-quality biopsy specimens or substantial missing clinical data were excluded. Eligible patients were approached and invited to participate in the study. A total of 257 patients were enrolled in this study. Oral informed consent was obtained from all participants before enrolment. The tool was piloted on a small subset of participants to assess clarity, feasibility and internal consistency. Minor modifications were made before final implementation. Patients were followed for a median duration of 20 months (3–38 months).

Study endpoint (outcome)

The primary endpoint (outcome) was relapse-free survival (RFS), the time from diagnosis to the first relapse event, including local recurrence, distant metastasis or death from any cause within 18 months.

Variables (selecting independent predictors of early relapse)

After an extensive literature review, core prognostic factors were identified [3, 13, 14, 25–30]. The set of variables used describes patient demographics, disease severity, known prognostic markers and treatment response. The final multivariable Cox proportional hazards model was not developed using an automated backward stepwise selection process; instead, a theory-driven and clinically informed variable selection approach was used. Candidate predictors were selected a priori based on established evidence from the literature, biological plausibility and their relevance to clinical decision-making in our setting. In addition, we considered:

- Clinical significance,
- Predictor interpretability,
- Data completeness,
- Avoidance of multicollinearity,
- Established prognostic relevance.

Variables with substantial missing data were excluded, and collinearity between candidate predictors was rigorously assessed before inclusion in the final model.

The selected predictors were age at diagnosis, lymph node (LN) status, tumour stage, immunohistochemical evaluation of human epidermal growth factor receptor 2 (HER2)] Ki-67, Topoisomerase 2A (TOP2A) and residual disease as a NACT response indicator. Certain factors, such as age and tumour stage, were initially assessed as both continuous and categorical variables. For consistency and interpretability, categorical variables were selected for the final model. Menopausal status was not considered due to its strong correlation with age, which is already included as a predictor. The categorical variables were defined using clinically relevant cut-off points, with each factor analysed relative to a predefined reference group. The reference groups are as follows: age (with less than 50 years as reference), tumour size (with less than 2 cm as reference), LN involvement (with negative LN as reference), stage (with stage I as reference), Ki-67 (less than 14% as reference), HER2 (negative as reference), Top2A (negative as reference) and complete response as a reference for treatment response.

Implementation

Patient data were extracted from medical records using a validated data collection form, which included variables such as clinical history, tumour characteristics, treatment regimen and follow-up outcomes. The data collection tool was administered to a small subset of participants to assess clarity, feasibility and internal consistency. Minor modifications were made before final implementation. The PI was constructed based on the regression coefficients (β) from the multivariable Cox regression model. A weighted scoring system was generated by standardising and scaling the coefficients, followed by rounding to produce an integer-based risk score. Patients were then classified into five risk groups based on their PI scores.

Statistical analysis

Statistical analysis was performed using SPSS software version 20.0 (SPSS, Chicago, IL). Descriptive analysis was performed. The primary endpoint was 18-month relapse (free survival). All statistical tests used for testing performance, prediction and discrimination abilities for the developed PI were carried out. Also, the validity of the index as a binary test was described by calculating the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The Discriminative ability of the newly developed PI was tested using several methods. First, the Kaplan–Meier test and two-sided log-rank test were employed to compare the five risk groups. Then, receiver operating characteristic (ROC) curves and area under the ROC curves (AUC) were used to describe the discrimination ability of the PI compared to the conventional NPI. Model calibration was carried out to describe how well the estimates of relapse from the model corresponded to the observed relapses using the Hosmer–Lemeshow test.

Analysis also included assessment of the reclassification ability of the new PI compared to the NPI. The net reclassification rate was calculated using the cross-table method and the equation below [31].

Net reclassification improvement (NRI) = *P* (up| event) −*P* (down| event) + *P* (down| non-event) −*P* (*up*| non-event)

P: Proportion of cases, up: in higher risk group, down: in lower risk group

Model validation was performed using bootstrapping with 1,000 resamples to assess internal validity and estimate optimism-corrected performance. Calibration was evaluated through the Hosmer–Lemeshow test and calibration plots, while discrimination was assessed using the C-index and ROC analysis.

Ethical consideration

The study was conducted in accordance with national research guidelines and the principles of the Declaration of Helsinki. Ethical approval was obtained from the IRB of the University of Khartoum and the Ministry of Health, Sudan, on 28 October 2018, under the chairmanship of Professor Sheikh Mahgoub. Written informed consent was waived by the IRB due to the observational nature of the study; however, oral informed consent was obtained from all participants before enrolment. Patient confidentiality was strictly maintained, and all data were anonymised.

Research

Results

Cohort characteristics

The study cohort comprised 257 patients, with a median age of 48 years (range: 24–70 years). As detailed in Table 1, the majority of patients were aged less than 50 years (59.1%). Pathological examination revealed a high rate of poorly differentiated disease (72%). LN involvement (86.4%), large tumour size (96.5%) and notable presentation of triple negative (TN) subtype. Most patients (82.5%) received. Anthracycline-taxane-based regimens were the most frequently used neoadjuvant regimen. Objective response was achieved in 64.5% of patients. About two-thirds of the study cohort experienced an early relapse event during the 2-year follow-up period of disease.

Establishment of the PI text for this sub-section

Results of the Cox Regression Model predicting 18-month relapse based on seven predictive factors, as illustrated in Table 2, revealed that of the seven predictors, initial LN involvement was the only statistically significant predictor (p = 0.004), contributing to a 3.9 times increase in the risk of relapse (HR = 3.79, 95% (CI:1.53–9.37). The tumour stage demonstrated a high effect size by contributing five times (HR = 5.18, 95% (CI:0.71–37.77); however, the association was not statistically significant (p = 0.105). Residual disease, HER2/neu and TOP2A contributed by >1 time despite not being statistically significant ($p > 0.05$). Ki-67 and age contributed less than <1 time, and their negative B coefficients indicate their reverse impact on this early relapse endpoint.

A weighted scoring system was generated based on the Wald z-score derived from the Cox regression coefficients, followed by scaling and rounding to produce an integer-based risk score. (Table 3). For example, for LN involvement, the wald z-score of 2.883 was multiplied by ten and rounded to 29, indicating a higher weight for this factor in predicting early relapse.

Table 1. Demographic, clinical, histopathological, treatment and response characteristics of the study cohort (*n* = 257).

Variable	Categories	*n*	%
Age in years (*n* = 257)	< 50 years	152	59.1
	≤50 years	105	40.9
Histopathological type (*n* = 257)	IDC	237	92.2
	Others	20	7.8
Tumor grade (*n* = 257)	Grade 1	2	0.8
	Grade 2	70	27.2
	Grade 3	185	72.0
Tumor differentiation (*n* = 257)	Poorly differentiated	185	72.0
	Well/moderately differentiated	72	28.0
Initial LN involvement (*n* = 257)	Involved	222	86.4
	Not involved	35	13.6
Initial tumor size (*n* = 257)	< 5 cm	9	3.5
	≥ 5 cm	248	96.5
Ki-67 (*n* = 75)	< 14%	10	13.3
	≥ 14%	65	86.7
HER2 (*n* = 252)	Positive	79	31.3
	Negative	173	68.7
Biological type (*n* = 257)	Luminal A	79	30.7
	Luminal B	64	24.9
	TN	75	29.2
	HER2 enriched	39	15.2
Top 2A expression (*n* = 130)	Positive	72	55.4
	Negative	58	44.6
Chemotherapy treatment (*n* = 257)	Anthracycline-based	45	17.5
	Taxane and Anthracycline-based	212	82.5
Overall good objective response (*n* = 257)	Responsive	166	64.6
	Non-responsive	91	35.6
Relapse /Recurrent disease	6-month relapse (refractory disease)	54	21.0
	24-month relapse (first 2 years relapse)	172	66.9

IDC = Invasive Ductal Carcinoma, HER2/neu = human epidermal growth factor receptor2 ,TN=Triple Negative, TOP2A = Topoisomerase II alpha

Table 2. The multivariable Cox regression hazard analysis of some demographic, clinical and treatment factors associated with early relapse in breast cancer patients (n = 257).

Variables	β	*p*-value	HR	95% CI (HR))	
				Lower	Upper
Age	−0.106	0.530	0.899	0.646	1.252
LN	1.332	0.004	3.788	1.531	9.372
Stage	1.644	0.110	5.178	0.710	37.769
HER2/neu	0.039	0.810	1.040	0.751	1.441
Ki-67	−0.018	0.470	0.982	0.935	1.031
TOP2A	0.016	0.470	1.016	0.973	1.060
Residual disease	0.140	0.670	1.151	0.605	2.190

β = Regression Coefficient, HR hazard ratio, LN Lymph Node involvement, HER2/neu = human epidermal growth factor receptor 2, TOP2A = Topoisomerase II alpha, statistical significance was set at $p < 0.05$

Table 3. Weighted risk score of the individual prognostic factors in the newly developed PI.

Predictors	β (Coefficient)	SE	Wald z-score (β/SE)	Assigned score
Age	−0.106	0.169	−0.627	−6
LN involvement	1.332	0.462	2.883	29
Tumor stage	1.644	1.014	1.621	16
HER2	0.039	0.166	0.234	2
Ki-67	−0.018	0.025	−0.720	−7
TOP2A	0.016	0.022	0.727	7
Residual disease	0.140	0.328	0.427	4
Theoretical range (−13–58)				

β = Regression Coefficient, SE = standard error, Wald z-score (β/SE) = weighted coefficient, LN Lymph Node involvement, HER2 = human epidermal growth factor receptor 2, TOP2A = Topoisomerase II alpha

Development of the risk group

The score was rescaled (divided by two) to enhance the clinical interpretability of risk categories. The categorisation was based on observed relapse rates, ensuring that each group had sufficient sample size and predictive accuracy. As shown in Table 4, the incidence of early relapses varied significantly across the risk groups, with relapse rates ranging from 0% in the very low-risk group to 84% in the very high-risk group.

Analytical characterisation of the PI

After bootstrapped internal validation, the model maintained high sensitivity (92.4%) and specificity (93.3%), an overall accuracy of 92.7%, indicating minimal overfitting and consistent predictive performance. The PI shows high accuracy in correctly identifying both patients who will relapse and those who will not. The PPV of 97.0% suggests that when the model predicts relapse, the prediction is highly likely to be correct. The NPV of 83.4% indicates that most patients classified as low risk by the model remained free from early relapse (Table 5).

Research

Table 4. Incidence of relapse events among the five risk groups according to the risk score calculated by the newly developed PI (n = 257).

Risk category	Patients (n)	Events (n)	Observed relapse %
Very low risk (group 1)	4	0	0.0%
Low risk (group 2)	18	2	11.0%
Moderate risk (group 3)	16	3	19.0%
High risk (group 4)	103	70	68.0%
Very high risk (group 5)	116	97	84.0%

Table 5. Diagnostic performance of the new developed PI.

Analytical characteristic	Value	95% CI
Sensitivity	92.4%	88.7%–95.9%
Specificity	93.3%	77.9%–99.2%
Positive likelihood ratio	13.860	3.600–52.900
Negative likelihood ratio	0.080	0.080–0.130
PPV	97.0%	89.4%–99.2%
NPV	83.4%	76.7%–89.3%
Accuracy	92.7%	88.7%–95.6%

Research

Discriminative ability

The Kaplan–Meier curves (Figure 1) demonstrate significantly different survival distributions across the five risk groups. The log-rank test confirmed that these differences were statistically significant ($p < 0.05$), suggesting that the PI can effectively stratify patients into risk categories with distinct RFS outcomes.

The ROC analysis (Figure 2) showed that the new PI had a higher discriminatory accuracy compared to the NPI, with an AUC of 0.81 compared to 0.74 for the NPI. An AUC closer to 1.0 indicates better predictive accuracy, and these results suggest that the new PI may provide more reliable prognostic information for early relapse in this population.

Calibration

The Hosmer–Lemeshow test (p = 1.00) indicated that the new PI demonstrated good calibration, meaning that the predicted relapse probabilities were closely aligned with the observed relapse events. This suggests that the PI is a reliable tool for predicting early relapses in the cohort.

Clinical risk reclassification

The NRI of 0.13 suggests that this new PI improved the predictive ability by 13%. The new PI tends to improve correct patient allocation in specific risk categories. The new PI improves the classification of patients compared to the conventional NPI risk classification by correctly reclassifying patients into more accurate risk categories based on their relapse risk. If externally validated patients at KOH are protected from being over- or undertreated. Its clinical value becomes more meaningful if the costs/benefits of correct/incorrect classifications are considered.

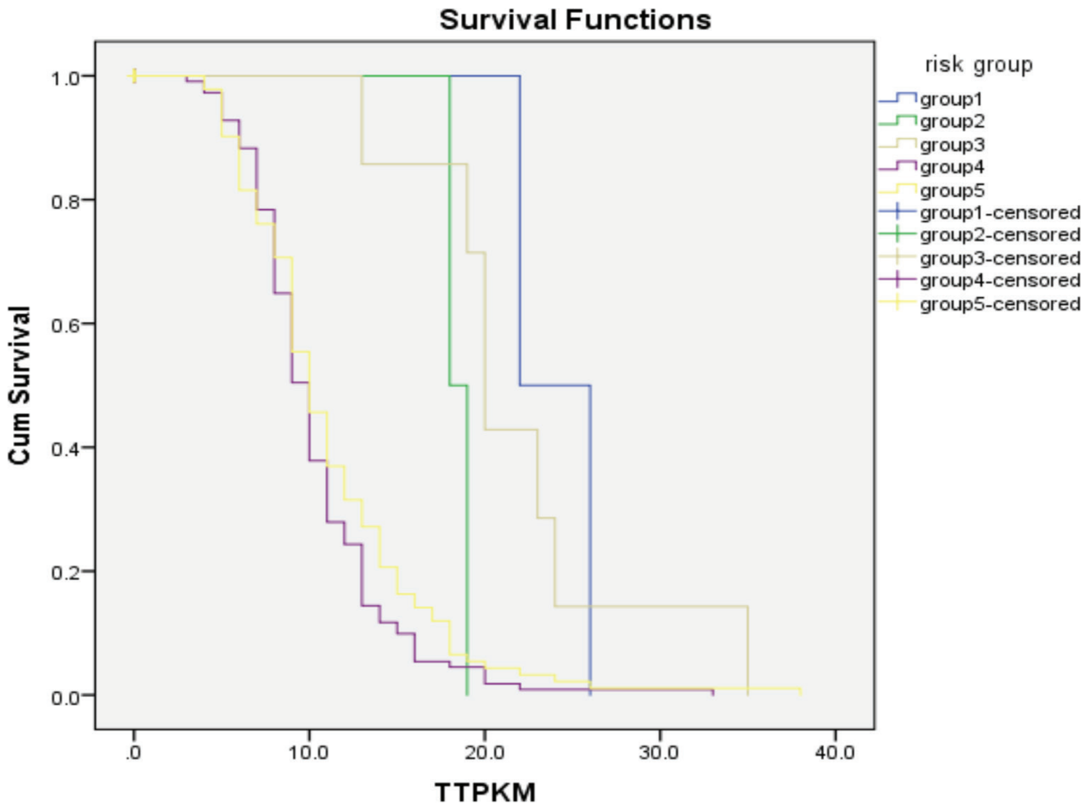

Figure 1. Kaplan–Meier curves of time to progression (TTP) in months of breast cancer patients at KOH stratified by risk groups. Kaplan–Meier analysis illustrating significantly different TTP probabilities across the five identified risk groups.

Research

Discussion

In this study, we successfully developed and internally validated a locally tailored PI for early relapse in breast cancer, showing superior discrimination to the NPI.

Previously published studies and guidelines guided the development process [11, 13, 19, 20, 22, 23].

Model predictors used in index development were consistent with other factors studied previously in comparable settings [3, 13, 19, 24, 25, 32–34].

While tools such as Oncotype DX and PAM50 have demonstrated significant prognostic value in high-income settings, their development cohorts are predominantly based on Western populations with different demographic and biological characteristics.

In contrast, breast cancer in Sudan often presents at a younger age and may exhibit distinct biological behaviour, including higher proliferative activity and different recurrence patterns.

Therefore, our locally developed model was designed to provide: context-specific risk stratification, reliance on readily available clinical variables, improved applicability in resource-limited settings and reduced dependence on costly genomic assays.

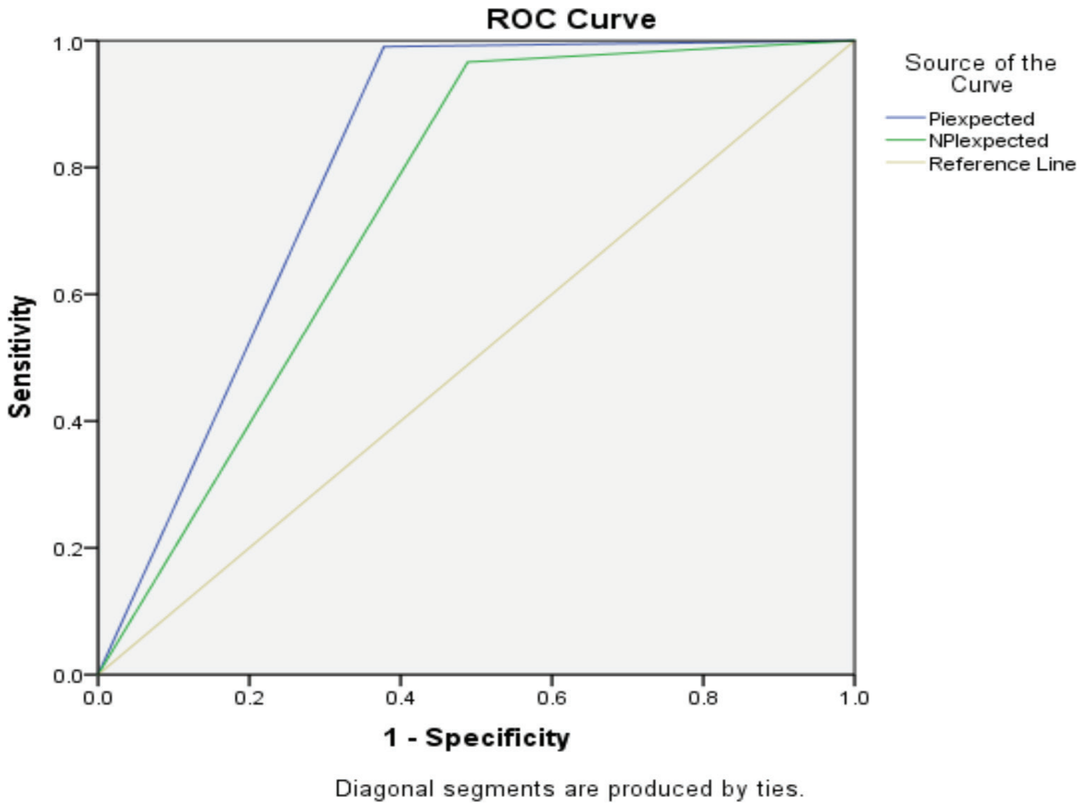

Figure 2. ROC curves comparing the predictive performance of the newly developed PI with the NPI. The ROC curve showing a superior discrimination with an AUC of 0.81 compared with 0.74 for the NPI.

The PI demonstrated promising analytical characteristics, including high sensitivity and specificity. The NPV of 83.4% indicates that most patients classified as low risk by the model remained free from early relapse, supporting the utility of the tool in identifying patients unlikely to experience recurrence. Together with the high PPV, this strengthens the model's value for clinical triage and resource allocation.

Future studies may further refine the model through the use of penalised regression techniques such as the Least Absolute Shrinkage and Selection Operator to improve variable selection, minimise overfitting and enhance model stability.

This revision better clarifies our modeling strategy and rationale. In addition, incorporating cost-effective inflammatory biomarkers, such as the neutrophil-to-lymphocyte ratio, to further improve predictive performance and external applicability should be considered.

This study was not without limitations. Awareness and consideration of many aspects that might affect the appropriate interpretation of findings help ensure reasonable PI performance. These aspects were related to study design, population heterogeneity, prognostic factors and analytical test selection. The weighted risk was derived from the Wald z-score(β/SE) of the multivariable cox model as a simple approach to reflect the relative contribution of each predictor. Given the proof-of-concept nature of the study the score should considered exploratory and interpreted with caution. The single-institute prospective design was used to improve the quality and decrease the heterogeneity in the development cohort. In addition, in spite of the lack of multiple imputation for missing biomarkers may have introduced minor bias, but internal bootstrapped validation confirmed stable coefficient estimates, supporting the model's reliability. Future multicenter validation is warranted to ensure external generalisability. The incorporation of machine learning approaches is also expected to enable the integration of multiple sources of information and help to identify the hidden patterns of prognostic factors in our local population. Finally, addressing the time of study conduction, Sudan has experienced significant disruption due to ongoing conflict, leading to further deterioration and fragmentation of the healthcare system. However, these challenges have not been accompanied by meaningful advancements in oncology decision-making practices or access to diagnostic and therapeutic resources. As such, the study findings remain relevant and reflective of current clinical realities in this setting.

Conclusion

The newly developed PI demonstrates the potential to significantly improve early relapse prediction in breast cancer patients at the KOH. This model may facilitate more individualised treatment decisions in high-risk patient and mitigate the risk of treatment-related morbidity in low-risk populations. The PI provides a practical framework to support personalised treatment decision in advanced breast cancer disease. Improving the model's performance is expected to pave the way for its clinical utility in the oncology setting in Sudan. Upon successful external validation, this PI offers a reliable, evidence-based metric tailored for resource-limited healthcare settings. The role of molecular and clinical pharmacists will be pivotal in incorporating genetic and molecular data into clinical practice, contributing to the advancement of personalised treatment strategies in the region.

List of abbreviations

CI, Confidence interval; NRI, Net reclassification improvement; PAM50, Prediction Analysis of Microarray 50; ROC-curve, Receiver operating characteristic curve; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement.

Acknowledgments

Special thanks go to all KOH staff, especially Histopathology, Statistics and the Medical Records unit. We also acknowledge the contribution and guidance from the oncologist in KOH, Dr Sulafa Ghurashi, Dr Mohaned, Dr Ibtihal and Dr Wala Karrory. Without whom this work would not have been possible.

Conflicts of interest

None.

Funding

No funds, grants or other support were received for this manuscript.

Author contributions

This work was carried out in collaboration among all authors. All authors were involved in concretisation, design and data interpretation. Authors MMA and NEOSH acquired data and took responsibility for the integrity of the data and data analysis. MMA, ME, MN and NEOSH performed the data analysis and interpretation of the results. KHM, BÍO, NEOSH and AAI contributed to the drafting and editing of the article and have granted their approval of the submitted manuscript. All authors read and approved the final manuscript.

References

1. Bray F, Laversanne M, and Sung H, *et al* (2024) **Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries** *CA Cancer J Clin* **74**(3) 229–263 [https://doi.org/10.3322/caac.21834] PMID: 38572751
2. Mohamed AA, Mohamed KEH, and Saad EA, *et al* (2018) **The response, operability, and type of surgery following neoadjuvant chemotherapy in Sudanese patients with locally advanced breast cancer** *Sudan J Med Sci (SJMS)* **13**(3) 219–229 [https://doi.org/10.18502/sjms.v13i3.2961]
3. Phung MT, Tin Tin S, and Elwood JM (2019) **Prognostic models for breast cancer: a systematic review** *BMC Cancer* **19** 230 https://doi.org/10.1186/s12885-019-5442-6 PMID: 30871490 PMCID: 6419427
4. Cardoso MJ, Poortmans P, and Senkus E, *et al* (2024) **Breast cancer highlights from 2023: knowledge to guide practice and future research** *Breast* **74** 103674 https://doi.org/10.1016/j.breast.2024.103674 PMID: 38340683 PMCID: 10869942
5. Trewin CB, Johansson ALV, and Hjerkind KV, *et al* (2020) **Stage-specific survival has improved for young breast cancer patients since 2000: but not equally** *Breast Cancer Res Treatment* **182** 477–489 https://doi.org/10.1007/s10549-020-05698-z
6. Martini R, Newman L, and Davis M (2022) **Breast cancer disparities in outcomes; unmasking biological determinants associated with racial and genetic diversity** *Clin Exp Metastasis* **39**(1) 7–14 https://doi.org/10.1007/s10585-021-10087-x PMCID: 9337912
7. Vanderpuye V, Grover S, and Hammad N, *et al* (2017) **An update on the management of breast cancer in Africa** *Infect Agents Cancer* **12** 13 https://doi.org/10.1186/s13027-017-0124-y
8. Muddather HF, Faggad A, and Elhassan MMA (2022) **Relapse-free survival in Sudanese women with non-metastatic breast cancer** *Glob Epidemiol* **8** 100082 https://doi.org/10.1016/j.gloepi.2022.100082
9. Guo H and Wei S (2025) **Prognostic models for breast cancer: a narrative review** *Hum Pathol* **162** 105835 https://doi.org/10.1016/j.humpath.2025.105835 PMID: 40513962
10. Riggio AI, Varley KE, and Welm AL (2021) **The lingering mysteries of metastatic recurrence in breast cancer** *Br J Cancer* **124** 13–26 https://doi.org/10.1038/s41416-020-01161-4 PMCID: 7782773
11. Sparano JA, Gray RJ, and Ravdin PM, *et al* (2019) **Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer** *N Engl J Med* **380**(25) 2395–2405 https://doi.org/10.1056/NEJMoa1904819 PMID: 31157962 PMCID: 6709671
12. Wang Y, Broeks A, and Giardiello D, *et al* (2023) **External validation and clinical utility assessment of PREDICT breast cancer prognostic model in young, systemic treatment-naïve women with node-negative breast cancer** *Eur J Cancer* **195** 113401 https://doi.org/10.1016/j.ejca.2023.113401
13. Khan O, Ajadi JO, and Almsned F, *et al* (2025) **Prognostic model for predicting recurrence in breast cancer patients in Saudi Arabia** *Sci Rep* **15** 18388 https://doi.org/10.1038/s41598-025-94530-z PMID: 40419677 PMCID: 12106668
14. Berton Giachetti PPM, Carnevale Schianca A, and Trapani D, *et al* (2025) **Current controversies in the use of Oncotype DX in early breast cancer** *Cancer Treat Rev* **135** 102887 https://doi.org/10.1016/j.ctrv.2025.102887 PMID: 40048856
15. Bhoo-Pathy N, Yip CH, and Hartman M, *et al* (2012) **Adjuvant! Online is overoptimistic in predicting survival of Asian breast cancer patients** *Eur J Cancer* **48**(7) 982–989 https://doi.org/10.1016/j.ejca.2012.01.034 PMID: 22366561
16. Nair NS, Kothari B, and Gupta S, *et al* (2023) **Validation of PREDICT Version 2.2 in a retrospective cohort of Indian women with operable breast cancer** *JCO Glob Oncol* **9** [https://doi.org/10.1200/GO.23.00114] PMID: 38085062 PMCID: 10846767

17. Nik Ab Kadir MN, Mohd Hairon S, and Ab Hadi IS, *et al* (2023) **A comparison between the online prognostic tool PREDICT and myBeST for women with breast cancer in Malaysia** *Cancers (Basel)* **15**(7) 2064 https://doi.org/10.3390/cancers15072064 PMID: 37046725 PMCID: 10093426

18. Collins GS, Reitsma JB, and Altman DG, *et al* (2015) **Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement** *BMJ* **350** g7594 https://doi.org/10.1136/bmj.g7594 PMID: 25569120

19. McLernon DJ, Giardiello D, and Van Calster B, *et al* (2022) **Assessing performance and clinical usefulness in prediction models with survival outcomes: practical guidance for Cox proportional hazards models** *medRxiv* 03.17.22272411 [https://doi.org/10.1101/2022.03.17.22272411]

20. Collins GS, Dhiman P, and Ma J, *et al* (2024) **Evaluation of clinical prediction models (part 1): from development to external validation** *BMJ* **384** 74819 https://doi.org/10.1136/bmj-2023-074819

21. Kent P, Cancelliere C, and Boyle E, *et al* (2020) **A conceptual framework for prognostic research** *BMC Med Res Methodol* **20**(1) 172 https://doi.org/10.1186/s12874-020-01050-7

22. Deliu N, Cottone F, and Collins GS, *et al* (2018) **Evaluating methodological quality of prognostic models including patient-reported health outcomes iN oncologY (EPIPHANY): a systematic review protocol** *BMJ Open* **8**(10) 25054https://doi.org/10.1136/bmjopen-2018-025054

23. Hueting TA, Van Maaren MC, and Hendriks MP, *et al* (2022) **The majority of 922 prediction models supporting breast cancer decision-making are at high risk of bias** *J Clin Epidemiol* **152** 238–247 https://doi.org/10.1016/j.jclinepi.2022.10.016

24. Mallett S, Royston P, and Dutton S, *et al* (2010) **Reporting methods in studies developing prognostic models in cancer: a review** *BMC Med* **8** 20 https://doi.org/10.1186/1741-7015-8-20 PMID: 20353578 PMCID: 2856521

25. He M, Zhang JX, and Jiang YZ, *et al* (2017) **The lymph node ratio as an independent prognostic factor for node-positive triple-negative breast cancer** *Oncotarget* **4**(27) 44870–44880 https://doi.org/10.18632/oncotarget.17413

26. Akbari ME, Akbari A, and Nafissi N, *et al* (2016) **Prognostic factors of recurrence (early and late) and death in breast cancer patients in Iranian women** *Int J Cancer Manag* **9**(6) 5747 [https://doi.org/10.17795/ijcp-5747]

27. Koca B, Kuru B, and Karabicak I, *et al* (2014) **Prognostic factors affecting disease-free survival in patients at age 35 or younger with invasive breast cancer** *Ann Ital Chir* **85**(3) 249–253 https://doi.org/https://dx.doi.org/85(3):249-53

28. Radmilović Varga L (2021) **Prognostic value of Topoisomerase 2-alpha and B-MYB in early breast cancer treated with adjuvant chemotherapy** *Acta Clin Croat* **60**(1) 16–24 [http://doi.org/10.20471/acc.2021.60.01.03]

29. Sánchez-Espinosa A, García-Rodríguez J, and Alonso-Aguirre V, *et al* (2025) **Expression of Fascin and DNA Topoisomerase 2-alpha in breast carcinoma: correlation with histological subtypes and other prognostic markers** *Int J Mol Sci* **26** 3076 https://doi.org/10.3390/ijms26073076 PMID: 40243780 PMCID: 11988776

30. Kerr KF, Wang Z, and Janes H, *et al* (2014) **Net reclassification indices for evaluating risk prediction instruments: a critical review** *Epidemiology* **25**(1) 114–121 https://doi.org/10.1097/EDE.0000000000000018

31. Qian F, Shen H, and Liu C, *et al* (2025) **Establishment and validation survival prediction models for T1 locally advanced breast cancer after breast conservation surgery versus mastectomy** *Sci Rep* **15**(1) 12189 https://doi.org/10.1038/s41598-025-91205-7 PMID: 40204827 PMCID: 11982249

32. Zhang M, Peng P, and Gu K, *et al* (2018) **Time-varying effects of prognostic factors associated with long-term survival in breast cancer** *Endocr Relat Cancer* **25**(5) 509–521 https://doi.org/10.1530/ERC-17-0502 PMID: 29472247 PMCID: 5862777

33. Ren J, Li Y, and Zhou J, *et al* (2024) **Developing machine learning models for personalized treatment strategies in early breast cancer patients undergoing neoadjuvant systemic therapy based on SEER database** *Sci Rep* **14** 22055 https://doi.org/https://doi.org/10.1038/s41598-024-72385-0

34. Ma T, Liu X-Y, and Cai S-L, *et al* (2024) **Development and validation of a nomogram for predicting rapid relapse in triple-negative breast cancer patients treated with neoadjuvant chemotherapy** *Front Cell Dev Biol* **12** 1417366 [https://doi.org/10.3389/fcell.2024.1417366]