Am J Transl Res 2026;18(5):4593-4604
www.ajtr.org /ISSN:1943-8141/AJTR0171985

Original Article

Development and validation of a nomogram for predicting grade 4 neutropenia in patients with breast cancer undergoing anthracycline-based chemotherapy

Jing Shi[1,2,3], Fuxing Zhao[2], Tianlei Qiu[2], Shifen Huang[2], Yongxin Li[2], Junli Ma[3], Jiuda Zhao[1,2]

[1]School of Clinical Medicine of Qinghai University, Affiliated Hospital of Qinghai University, Xining 810000, Qinghai, China; [2]Breast Disease Diagnosis and Treatment Center, Affiliated Hospital of Qinghai University and Affiliated Cancer Hospital of Qinghai University, Xining 810000, Qinghai, China; [3]Affiliated Hospital of Jining Medical University, Jining 272029, Shandong, China

Received February 9, 2026; Accepted May 18, 2026; Epub May 25, 2026; Published May 30, 2026

Abstract: Background: Grade 4 neutropenia is a severe toxicity associated with anthracycline-based chemotherapy in patients with breast cancer. This study aimed to develop and validate a clinical prediction model for this adverse event during the first cycle of anthracycline-based chemotherapy. Methods: In this multicenter retrospective study, female patients with breast cancer who received their first cycle of anthracycline-based chemotherapy between January 2017 and December 2023 were included. Patients were randomly assigned into training and internal validation cohorts in a 2:1 ratio. An independent external validation cohort comprised patients treated between January 2024 and December 2025. Predictors were selected using multivariable logistic regression with a backward stepwise approach, and a nomogram was constructed. Model performance was assessed in terms of discrimination, calibration, and clinical utility using decision curve analysis (DCA). Results: A total of 672 patients were analyzed, with 518 in the training and internal validation cohort and 154 in the external validation cohort. Four independent predictors of grade 4 neutropenia were identified: age, body mass index (BMI), Ki67 index, and chemotherapy regimen. The model achieved area under the curve (AUC) values of 0.757, 0.759, and 0.731 in the training, internal validation, and external validation cohorts, respectively, demonstrating good discriminative ability. Calibration plots demonstrated satisfactory agreement between predicted and observed risks across all cohorts, and DCA confirmed the clinical utility of the model. Conclusion: We developed and externally validated a robust nomogram for predicting grade 4 neutropenia in patients with breast cancer undergoing their first cycle of anthracycline-based chemotherapy. This user-friendly tool may assist clinicians in identifying high-risk patients prior to treatment initiation, enabling timely preventive intervention.

Keywords: Anthracycline, breast cancer, grade 4 neutropenia, nomogram, prediction model

Introduction

Breast cancer is the most commonly diagnosed non-skin cancer and the leading cause of cancer-related death among women worldwide [1-3]. Chemotherapy remains a cornerstone for breast cancer treatment, and anthracycline-based regimens, including those containing epirubicin and pirarubicin, are widely utilized due to their efficacy in improving survival [4-6]. However, the therapeutic gains of these regimens are usually limited by substantial treatment-related toxicities.

Neutropenia, a common consequence of chemotherapy-induced myelosuppression, frequently necessitates dose reductions or delays in chemotherapy, which may compromise treatment efficacy [7]. The incidence of febrile neutropenia can be up to 50% in patients with grade 4 neutropenia (absolute neutrophil count [ANC] < 0.5×10^9/L), with an associated mortality rate of approximately 5-10% [8]. Grade 4 neutropenia is a severe side effect of chemotherapy, that can lead to treatment interruptions and even mortality [9, 10]. Previous studies have reported associations between neu-

https://doi.org/10.62347/XUPB3201

tropenia and several factors such as patient age, sex, chemotherapy regimens, and comorbidities (e.g., diabetes and chronic renal disease) [11, 12].

Granulocyte colony-stimulating factor (G-CSF) is commonly used for primary prophylaxis against chemotherapy-induced neutropenia. Two main formulations are available: short-acting G-CSF (e.g., filgrastim), which requires daily injections for several days and is relatively expensive, and long-acting pegylated G-CSF (e.g., pegfilgrastim), which is administered as a single injection per chemotherapy cycle, offering greater convenience but at a significantly higher cost [13]. This cost difference may impose a substantial economic burden on patients and healthcare systems, particularly in resource-limited settings. In addition, while G-CSF effectively reduces the risk of infection, it is associated with adverse events such as bone pain and fever [14, 15]. Therefore, early detection of and risk management are critical to optimizing therapeutic strategies and improving outcomes for patients undergoing chemotherapy.

Most existing models for predicting chemotherapy-induced neutropenia have been developed and validated in patients with cancers other than breast cancer, or in patients receiving non-anthracycline-based treatment, which may limit their applicability in breast cancer patients initiating anthracycline-based chemotherapy [16-18]. As a result, there remains a significant unmet need for a disease-specific tool that incorporates readily available pretreatment factors to enable accurate, individualized risk assessment prior to the first chemotherapy cycle.

The present study seeks to identify independent predictors for grade 4 neutropenia following the first cycle of anthracycline-based chemotherapy in patients with breast cancer using retrospective methods. Based on these predictors, we sought to establish a nomogram using routinely collected clinical and pathologic data to estimate individualized risk, enabling clinicians to implement optimized prophylactic strategies. Accurate predictive models are essential for personalizing therapy, reducing side effects, and improving treatment outcomes. Ultimately, this study may enhance clinical decision-making and patient management for breast cancer treatment.

Patients and methods

Study population and data collection

This multicenter retrospective study included female patients with pathologically confirmed primary breast cancer who underwent curative surgery followed by their first cycle of anthracycline-based chemotherapy between January 2017 and December 2023. Chemotherapy regimens consisted of either epirubicin-based (EC: epirubicin 90-100 mg/m^2, cyclophosphamide 600 mg/m^2) or pirarubicin-based (AC: pirarubicin 50 mg/m^2, cyclophosphamide 600 mg/m^2) protocols.

The inclusion criteria were as follows: (1) Complete clinical data; (2) Female patients with an age of ≥ 18 years; (3) Completion of the first cycle of the target chemotherapy regimen; (4) No prior history of chemotherapy or radiotherapy; and (5) Status post curative-intent operation for breast cancer. Pre-chemotherapy hematologic requirements were: hemoglobin ≥ 95 g/L, ANC ≥ 1.8 × 10^9/L, and platelet count ≥ 80 × 10^9/L. Patient with incomplete medical records or those receiving long-acting G-CSF were excluded.

In this study, short-acting prophylactic G-CSF was defined as the recombinant human G-CSF (rhG-CSF) administered at a dose of 5 μg/kg/day for 2-3 consecutive days, initiated 24 hours after completion of chemotherapy, as primary prophylaxis based on physician's discretion. rhG-CSF has a relatively short duration of action, requiring daily injections, but is substantially less expensive and widely accessible in clinical practice. Long-acting prophylactic G-CSF was defined as pegylated G-CSF (6 mg) administered once 24 hours after completion of chemotherapy, providing sustained hematopoietic support throughout the chemotherapy cycle with greater convenience, albeit at higher cost. Data on G-CSF administration, including drug name, dosage, start date, and duration, were extracted from the electronic medical records.

Clinicopathologic, demographic, and treatment information was retrieved from the electronic medical records of Qinghai University Affiliated

Hospital and Jining Medical University Affiliated Hospital. Variables analyzed included patient age, body mass index (BMI), hormone receptor (HR) status, Ki67 index, human epidermal growth factor receptor 2 (HER2) status, histologic type, T stage, N stage and pathologic stage, chemotherapy regimen, and use of short-acting G-CSF as prophylaxis. Complete blood counts were collected between the end of the first chemotherapy cycle and the start of the second cycle.

Patients recruited between 2017 and 2023 were randomly divided at a 2:1 ratio into a training cohort and an internal validation cohort. An independent external validation cohort consisted of patients meeting the same criteria treated at the same hospitals between January 2024 and December 2025. This retrospective study was approved by the Ethics Committee of Qinghai University Affiliated Hospital. The protocol was conducted in accordance with the ethical standards of the Declaration of Helsinki, and written informed consent was waived by the Ethics Committee due to the retrospective nature of the study.

Outcome definition

The primary outcome of this study was grade 4 neutropenia (ANC < 0.5×10^9/L) according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 5.0 [19]. Blood counts were assessed using conventional laboratory testing within 21 days after completion of the first chemotherapy cycle. The lowest ANC value observed within this 21-day window was used for analysis.

Statistical analysis

Sample size estimation: Based on the events per variable (EPV) rule of thumb, which recommends at least 10 events per candidate predictor to ensure reliable logistic regression estimates, we anticipated including up to 5 candidate variables in the final multivariable model. Thus, a minimum of 50 events (grade 4 neutropenia cases) was required. Considering an estimated incidence of grade 4 neutropenia of approximately 30-35% in our population, we planned to enroll at least 150-170 patients in the training cohort. The final training cohort (n = 346 with 115 events) exceeded this minimum requirement.

Sample size adequacy: A *post-hoc* sample size adequacy analysis was conducted using the EPV criterion to confirm model robustness [20, 21]. In our training cohort (n = 346), there were 115 events, yielding an EPV of approximately 23, which exceeds the commonly recommended minimum of 10, confirming adequacy for stable logistic regression estimates.

Cohort assignment: Patients treated between 2017 and 2023 were randomly assigned at a 2:1 ratio to the training and internal validation cohort using computer-generated random number. An independent external validation cohort included patients meeting the same inclusion criteria at the same institutions between January 2024 and December 2025.

Predictor selection and model development: Candidate variables were identified using univariate logistic regression analysis ($P < 0.1$). Clinically relevant variables (e.g., preventive therapy) were pre-specified. A multivariable logistic regression model was developed using backward stepwise selection. The findings are expressed as odds ratios (ORs) with 95% confidence intervals (CIs).

Nomogram construction: A nomogram was constructed according to the final multivariable model to estimate individualized risk of grade 4 neutropenia.

Model validation: Model performance was assessed in all three cohorts. Discrimination was measured using the area under the receiver operating characteristic curve (AUC). Calibration was assessed visually using calibration plots and statistically using the Hosmer-Lemeshow goodness-of-fit test. All cohorts were bootstrapped (n = 200 replicates) to generate bias-corrected calibration curves. Clinical applicability was assessed using decision curve analysis (DCA) to calculate the net clinical benefit at various threshold probabilities.

Statistical software: All statistical analyses were performed using R software (version 4.3.3; R Foundation for Statistical Computing; Vienna, Austria).

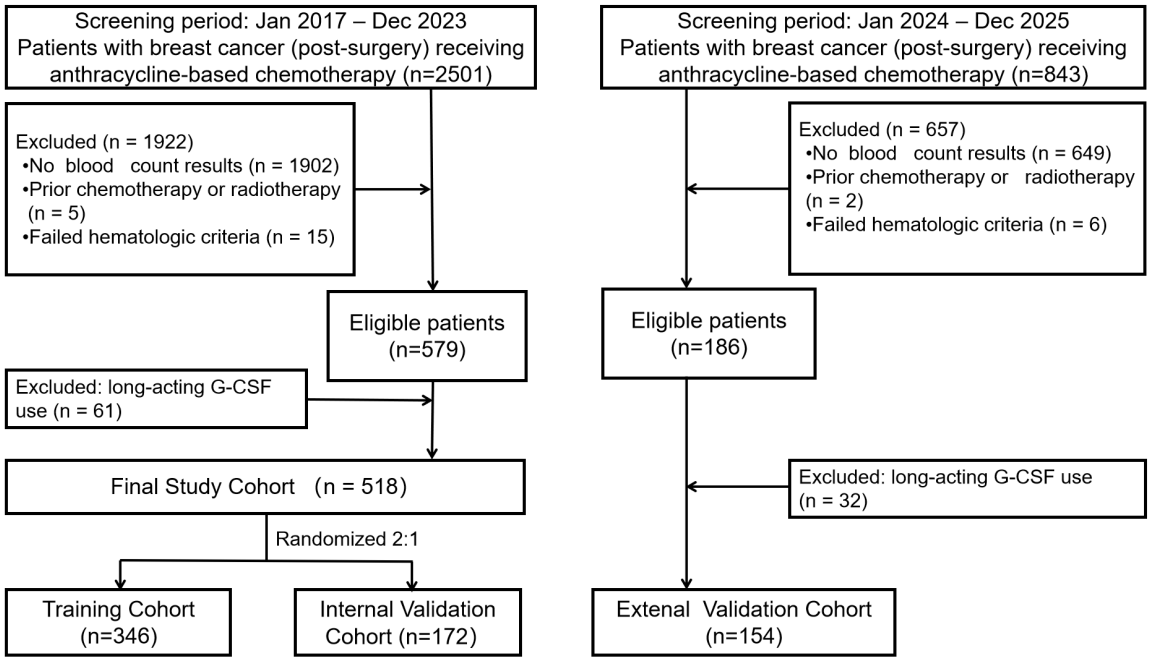

Figure 1. Flow diagram of patient selection and cohort allocation. G-CSF, granulocyte colony-stimulating factor.

For comparisons among the three cohorts, categorical variables were analyzed using the chi-square test or Fisher's exact test, as appropriate. A two-sided $P < 0.05$ was considered significant.

Results

Patient characteristics

A total of 2,501 patients with breast cancer were screened between January 2017 and December 2023. After applying the pre-specified inclusion and exclusion criteria, the final analytical population consisted of 518 patients. These patients were randomly allocated into two cohorts for model development and internal validation: 346 patients in the training cohort and 172 patients in the internal validation cohort. The participant selection process is presented in **Figure 1**. Grade 4 neutropenia occurred in 173 patients (33.4%), with similar incidence in the training (n = 115, 33.2%) and internal validation (n = 58, 33.7%) cohorts ($P > 0.05$).

An independent external validation cohort was formed by enrolling patients meeting the same inclusion criteria at the same institutions between January 2024 and December 2025, comprising 154 eligible patients. The incidence of grade 4 neutropenia in the external cohort was 32.5% (50/154).

As shown in **Table 1**, baseline characteristics were well balanced across the three cohorts (all $P > 0.05$), confirming the comparability of the datasets.

Predictor selection

Univariate logistic regression analysis identified four variables significantly associated with grade 4 neutropenia ($P < 0.05$), including age, BMI, Ki67 index, and chemotherapy regimen. Preventive treatment, which was not statistically significant in univariable analyses ($P = 0.19$), was pre-specified for inclusion in the multivariable model due to its clinical relevance in decision-making.

The final multivariable model retained four strong independent predictors and included all five candidate variables (**Table 2**). Age ≥ 50 years was associated with an increased risk of grade 4 neutropenia (OR = 1.94; 95% CI: 1.18-3.19, $P = 0.01$). Conversely, BMI ≥ 25 kg/m^2 (OR = 0.48; 95% CI: 0.29-0.79, $P = 0.004$), Ki67 index ≥ 20% (OR = 0.49; 95% CI: 0.30-0.81, $P = 0.005$), and AC regimen of chemo-

Table 1. Comparison of baseline characteristics among cohorts

Variable	Training (n = 346)	Internal Validation (n = 172)	External Validation (n = 154)	χ^2	*p*
Age				4.56	0.10
< 50	168 (48.55%)	95 (55.23%)	67 (43.51%)		
≥ 50	178 (51.45%)	77 (44.77%)	87 (56.49%)		
BMI				0.59	0.74
< 25	189 (54.62%)	100 (58.14%)	85 (55.19%)		
≥ 25	157 (45.38%)	72 (41.86%)	69 (44.81%)		
HR status				1.40	0.50
Negative	103 (29.79%)	60 (34.88%)	48 (31.17%)		
Positive	243 (70.23%)	112 (65.12%)	106 (68.83%)		
Ki67 index				0.32	0.85
< 20	141 (40.75%)	67 (38.95%)	59 (38.31%)		
≥ 20	205 (59.25%)	105 (61.05%)	95 (61.69%)		
HER2 status				2.54	0.28
Negative	249 (71.97%)	123 (71.51%)	102 (66.23%)		
Positive	84 (24.38%)	40 (23.26%)	47 (30.52%)		
Histologic type				4.79	0.09
IDC	307 (88.73%)	148 (86.05%)	144 (93.51%)		
Other	39 (11.27%)	24 (13.95%)	10 (6.49%)		
T stage				1.48	0.48
T0-1	123 (35.55%)	52 (30.23%)	51 (33.12%)		
T2-4	223 (64.45%)	120 (69.77%)	103 (66.88%)		
N stage				1.50	0.83
N0	128 (36.99%)	61 (35.47%)	53 (34.42%)		
N1	135 (39.02%)	63 (36.63%)	64 (41.56%)		
N2-3	83 (23.99%)	48 (27.91%)	37 (24.03%)		
Pathologic stage				1.15	0.89
Stage I	52 (15.03%)	27 (15.70%)	21 (13.64%)		
Stage II	208 (60.12%)	97 (56.40%)	95 (61.69%)		
Stage III	86 (24.86%)	48 (27.91%)	38 (24.68%)		
Chemotherapy regimen				1.09	0.58
EC	142 (41.04%)	73 (42.44%)	57 (37.01%)		
AC	204 (58.86%)	99 (57.56%)	97 (62.99%)		
Preventive therapy				0.58	0.75
No	121 (34.97%)	55 (31.98%)	50 (32.47%)		
Yes	225 (65.04%)	117 (68.02%)	104 (67.53%)		

Notes: IDC: invasive ductal carcinoma; BMI: body mass index; HR: hormone receptor; HER2: human epidermal growth factor receptor 2; EC: epirubicin plus cyclophosphamide; AC: pirarubicin plus cyclophosphamide. *p*-values were calculated using chi-square test for comparisons among the three cohorts. HER2 status was unavailable for 27 patients and is reported as missing. These cases were excluded from the denominator for HER2-related analyses.

therapy (OR = 0.23; 95% CI: 0.14-0.38, $P < 0.001$) were identified as independent protective factors. Preventive therapy showed a trend toward increased risk (OR = 1.56, 95% CI: 0.92-2.64), although this association was not significant in the final multivariable model ($P = 0.10$).

Nomogram development and clinical application

To optimize clinical application, a visual predictive tool (nomogram) was constructed based on the four significant predictors identified in the final multivariable regression analysis,

Table 2. Univariable and multivariable logistic regression analyses of factors associated with grade 4 neutropenia in the training cohort

Variable	Total (n)	Univariate analysis		Multivariate analysis	
		OR (95% CI)	*p* value	OR (95% CI)	*p* value
Age					
< 50	168	Reference		Reference	
≥ 50	178	1.72 (1.09-2.71)	0.02	1.94 (1.18-3.19)	0.01
BMI					
< 25	189	Reference		Reference	
≥ 25	157	0.54 (0.34-0.85)	0.01	0.48 (0.29-0.79)	0.004
HR status					
Negative	103	Reference			
Positive	243	0.86 (0.53-1.39)	0.54		
Ki67 index					
< 20	141	Reference		Reference	
≥ 20	205	0.48 (0.31-0.76)	0.002	0.49 (0.30-0.81)	0.005
HER2 status					
Negative	249	Reference			
Positive	84	1.00 (0.59-1.69)	1.00		
Histological type					
IDC	307	Reference			
Other	39	1.13 (0.56-2.26)	0.74		
T stage					
T0-1	123	Reference			
T2-4	223	1.28 (0.80-2.05)	0.31		
N stage					
N0	128	Reference			
N1	135	1.18 (0.70-1.98)	0.53		
N2-3	83	1.36 (0.76-2.44)	0.30		
Pathologic stage					
Stage I	52	Reference			
Stage II	208	1.52 (0.76-3.04)	0.23		
Stage III	86	1.87 (0.87-4.01)	0.11		
Chemotherapy regimen					
EC	142	Reference		Reference	
AC	204	0.23 (0.14-0.36)	< 0.001	0.23 (0.14-0.38)	< 0.001
Preventive therapy					
No	121	Reference		Reference	
Yes	225	1.38 (0.86-2.23)	0.19	1.56 (0.92-2.64)	0.10

Notes: IDC: invasive ductal carcinoma; BMI: body mass index; HR: hormone receptor; HER2: human epidermal growth factor receptor 2; EC: epirubicin plus cyclophosphamide; AC: pirarubicin plus cyclophosphamide. Variables with $P < 0.1$ in univariable analysis (age, BMI, Ki67 index, chemotherapy regimen) and preventive therapy (pre-specified based on clinical relevance) were included in multivariable logistic regression with backward stepwise selection.

including age, BMI, Ki67 index, and chemotherapy regimen (**Figure 2**). In addition, preventive therapy was also incorporated into the nomogram despite not being an independent significant predictor. By incorporating this variable in the model accordingly, clinicians can time-dependently estimate an individual patient's risk of grade 4 neutropenia under two conditions: (1) The natural course without intervention, and (2) Accounting for the effect of preventive therapy. This nomogram enables clinicians to directly calculate an individualized risk

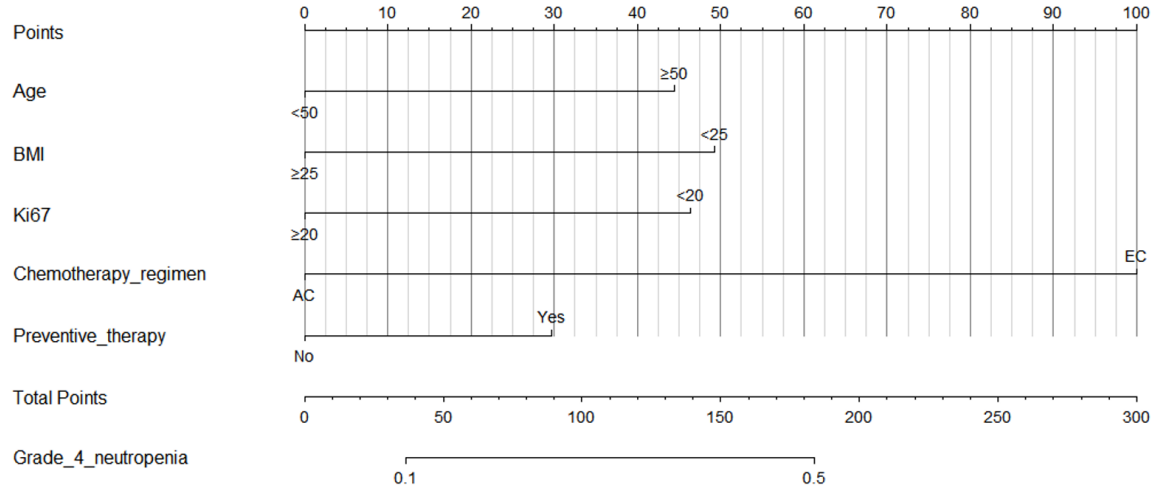

Figure 2. Nomogram for predicting the risk of grade 4 neutropenia. Notes: EC, epirubicin plus cyclophosphamide; AC, pirarubicin plus cyclophosphamide.

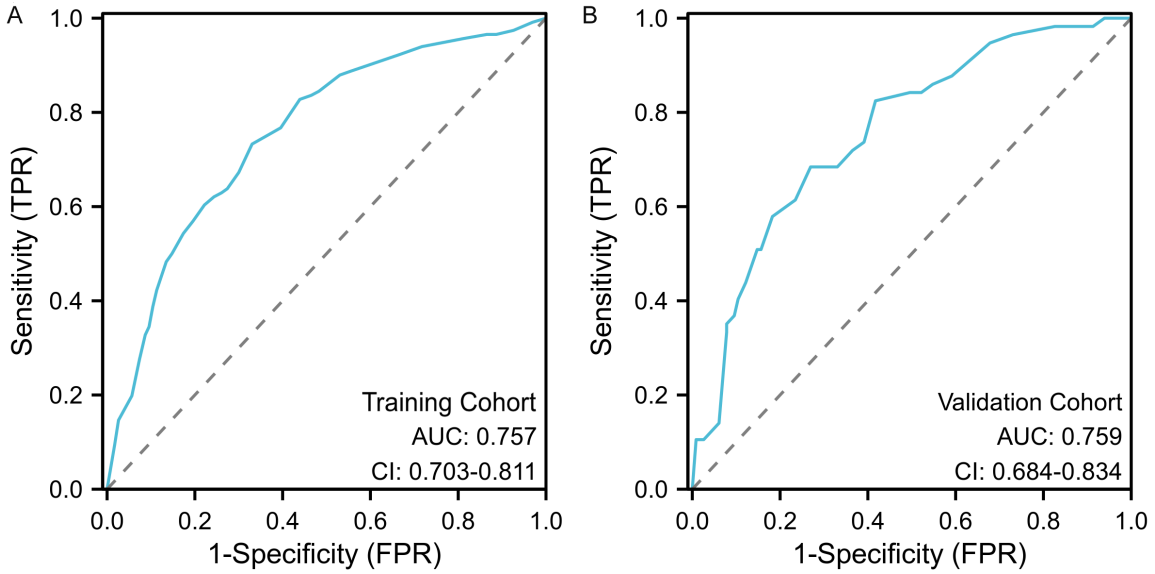

Figure 3. Receiver operating characteristic (ROC) curves of the nomogram in the training and internal validation cohorts. A. Training cohort, area under the curve (AUC) = 0.757; B. Internal validation cohort, AUC = 0.759.

score, facilitating bedside risk stratification and personalized management.

Model performance and internal validation

In the training cohort, the prediction model demonstrated good discriminative power with an AUC of 0.757 (95% CI: 0.703-0.811; **Figure 3A**). This performance was reproducible in the internal validation cohort, with a model AUC of 0.759 (95% CI: 0.684-0.834; **Figure 3B**).

Moreover, calibration was excellent in both cohorts, as supported by Hosmer-Lemeshow goodness-of-fit tests, which showed no significant deviation between predicted and observed events rates in either the training cohort (P = 0.700) or internal validation cohort (P = 0.191). Consistent calibration across these datasets, visually confirmed by calibration curves (**Figure 4A**, **4B**), suggested that the model provides accurate risk prediction without evidence of overfitting. This high degree of repro-

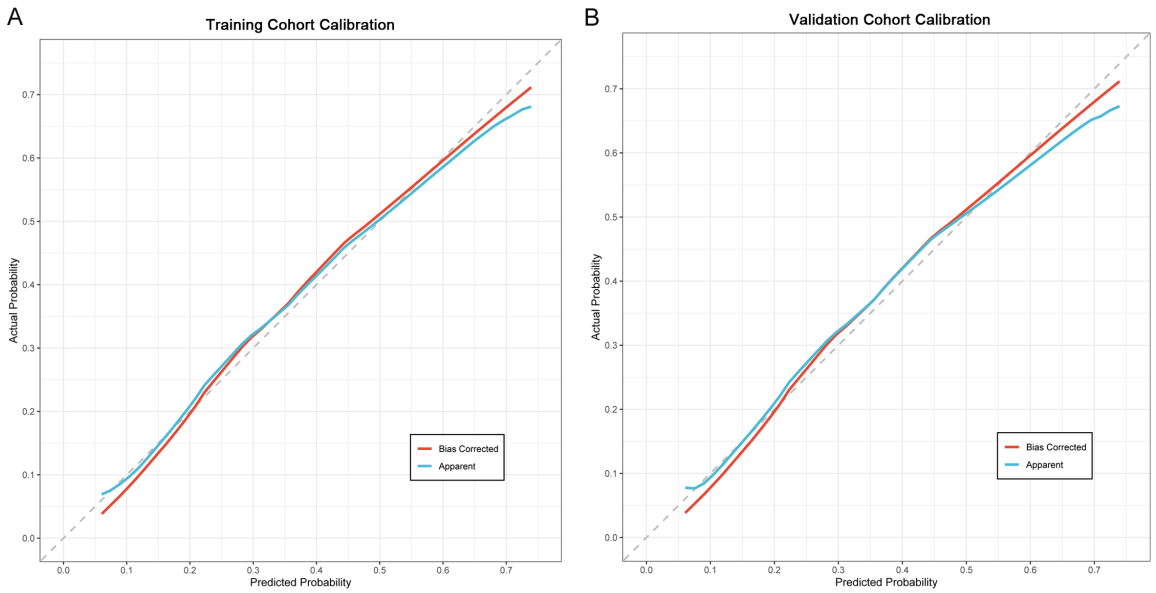

Figure 4. Calibration curves of the nomogram in the training and internal validation cohorts. A. Training cohort, Hosmer-Lemeshow test *P* = 0.700; B. Internal validation cohort, Hosmer-Lemeshow test *P* = 0.191.

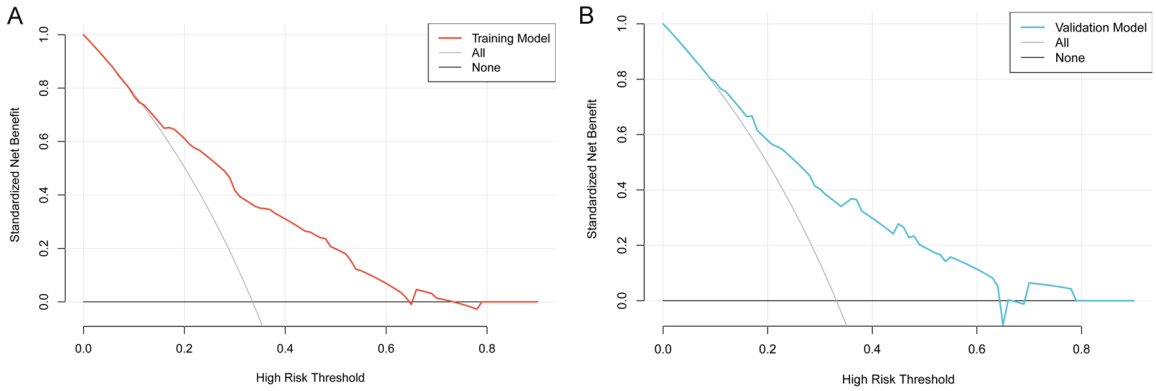

Figure 5. Decision curve analysis (DCA) of the nomogram in the training and internal validation cohorts. A. Training cohort; B. Internal validation cohort.

ducibility indicates the model is stable and reliable for clinical application.

Clinical use

DCA was performed to estimate the clinical utility of the nomogram in the training and internal validation cohorts. As shown in **Figure 5**, the nomogram produced greater net benefit compared with the "treat-all" or "treat-none" strategies over most of clinically relevant threshold probabilities. Within the 20%-60% threshold probability range, the model consistently demonstrated positive net benefit, confirming its utility in identifying suitable candidates for prophylactic intervention against grade 4 neutropenia.

External validation

The generalizability of the model was evaluated in an independent external cohort comprising 154 patients treated between January 2024 and December 2025. Baseline data of the external validation cohort are presented in **Table 1**. When applied to this validation cohort, the model demonstrated good discriminatory performance with an AUC of 0.731 (95% CI: 0.648-0.815; **Figure 6A**). Calibration was also satisfactory, as visually confirmed by calibra-

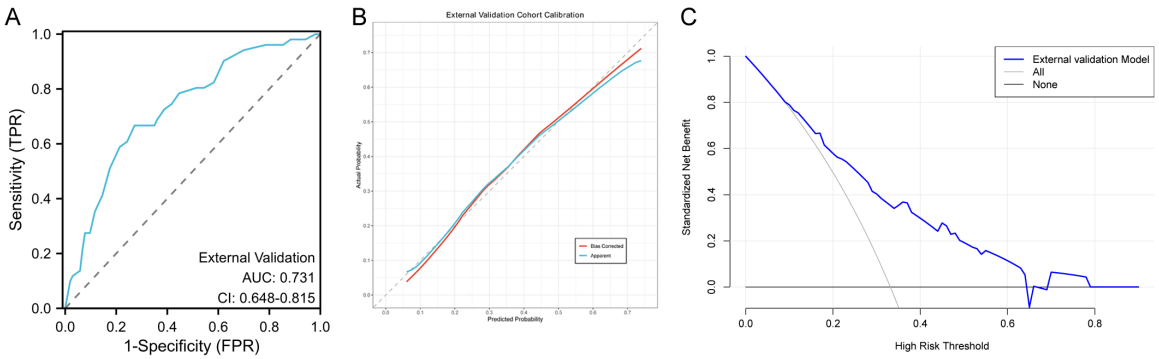

Figure 6. Performance of the nomogram in the external validation cohort. A. Receiver operating characteristic (ROC) curve in the external validation cohort, AUC = 0.731; B. Calibration curve in the external validation cohort (Hosmer-Lemeshow test P = 0.625); C. Decision curve analysis (DCA) in the external validation cohort.

tion plot (**Figure 6B**) and statistically verified by a non-significant Hosmer-Lemeshow goodness of fit test value (P = 0.625). DCA also suggested that the nomogram provided a positive net benefit across a wide range of threshold probabilities (**Figure 6C**), supporting its potential applicability in clinical practice.

Discussion

In this multicenter study, we developed and validated a clinical model to predict the occurrence of grade 4 neutropenia after the first cycle of anthracycline-based chemotherapy in patients with breast cancer. Four independent predictors were identified: age, BMI, Ki67 index, and chemotherapy regimen. The model incorporating these predictors along with the clinically relevant variable of preventive therapy demonstrated good and stable discriminative performance, with AUCs of 0.757 in the training cohort, 0.759 in the internal validation cohort, and 0.731 in the temporal external validation cohort. Calibration was excellent across all cohorts and DCA confirmed a net clinical benefit over a broad range of probability thresholds. The final model was transformed into a clinically useful nomogram, providing a practical tool for individualized risk estimation.

Our findings are consistent with, and expand upon current evidence on risk factors for chemotherapy-induced toxicity. Older age (≥ 50 years) was associated with a higher risk of grade 4 neutropenia, reflecting age-related declines in organ function and hematopoietic reserve, which may impair drug clearance and bone marrow recovery [22, 23]. In contrast, higher BMI (≥ 25 kg/m^2) emerged as a protective factor. This observation may be attributed to the fact that chemotherapy dosing is often based on body surface area, and pharmacokinetics may not be independent of body fat percentage. Altered drug distribution and metabolism in individuals with higher adiposity could lead to relative under-treatment and consequently lower hematologic toxicity [24]. Similar findings have been reported in colon cancer patients, where low BMI was identified as an independent risk factor for neutropenia [25].

Ki67 index is closely associated with the tumor cell proliferation and is frequently used to evaluate chemotherapy response and predict prognosis [26, 27]. Patients with Ki67 ≥ 20% had a significantly lower incidence of grade 4 neutropenia, a counterintuitive finding that has rarely been reported. One possible explanation is that tumors with high proliferative activity may be more sensitive to anthracycline-based chemotherapy, leading to rapid tumor cell kill and potentially a lower sustained burden on the bone marrow microenvironment. Alternatively, this observation may reflect an epiphenomenon associated with other biological features of aggressive tumor subtypes. Further prospective studies are warranted to elucidate the underlying mechanisms.

The chemotherapy regimen was also a strong predictor, with the AC regimen associated with a substantially lower risk of grade 4 neutropenia compared with the EC regimen. This difference may be attributable to variations in

drug potency, pharmacokinetics, or myelosuppressive profiles between the two anthracyclines, although direct comparative studies are limited.

Although HR status is a fundamental biomarker in breast cancer for guiding prognosis and treatment strategies [28], it did not correlate with neutropenia risk in our analysis. This finding is notable, as experimental data suggest that estrogen signaling can influence both tumor growth and hematopoietic stem cell differentiation [29, 30]. The absence of a protective effect in HR-positive patients may indicate that the potent cytotoxic effects of anthracycline-based chemotherapy override any protective signals from endogenous estrogen in the bone marrow microenvironment. Alternatively, tumor HR status may not accurately reflect systemic hormone levels or estrogen signaling capacity within the bone marrow.

The greatest novelty of the present study lies in the construction and detailed, multi-step validation of the first dedicated nomogram to predict grade 4 neutropenia after first-cycle of anthracycline-based chemotherapy in breast cancer. Successful external validation (AUC 0.731) supports its generalizability to real-world settings. Notably, the model achieved robust performance using only routinely available pre-treatment variables, enhancing its practicality and feasibility for clinical implementation. From a clinical perspective, the nomogram enables clinicians to stratify patients into high- and low-risk categories prior to chemotherapy initiation, allowing for more informed decisions regarding prophylactic G-CSF use.

This study has several limitations. First, its retrospective design may have introduced bias from unmeasured confounders. Second, even though the inclusion of an independent external validation cohort provided evidence of generalizability, further validation in more diverse and geographically diverse populations is warranted. Third, the biological significance of Ki-67 index in predicting neutropenia requires further experimental confirmation. Future studies should focus on larger, multi-institutional, prospective trials to further validate these findings while minimizing the biases inherent in retrospective analyses.

Conclusion

We developed and internally validated a clinically relevant prediction model for grade 4 neutropenia based on routinely available pre-treatment variables. The model demonstrated excellent reproducibility and predictive performance. The nomogram provides an efficient and straightforward way to shift from a "one-size-fits-all" approach toward individualized risk-adapted management of neutropenia, which is essential for optimizing the safety and efficacy of chemotherapy in patients with breast cancer. Moreover, successful external validation supports its generalizability and readiness for clinical implementation.

Acknowledgements

This work was supported by grants from the National Nature Science Foundation of China (No. 82160859).

Disclosure of conflict of interest

None.

Address correspondence to: Jiuda Zhao, Breast Disease Diagnosis and Treatment Center, Affiliated Hospital of Qinghai University and Affiliated Cancer Hospital of Qinghai University, No. 29, Tongren Road, Xining 810000, Shandong, China. Tel: +86-9716230-893; E-mail: jiudazhao@126.com

References

[1] DeSantis C, Siegel R, Bandi P and Jemal A. Breast cancer statistics, 2011. CA Cancer J Clin 2011; 61: 409-418.

[2] Xu H and Xu B. Breast cancer: epidemiology, risk factors and screening. Chin J Cancer Res 2023; 35: 565-583.

[3] Wang J, Wang S, Huang E and Ma F. Expert consensus on tiered diagnosis and treatment of breast cancer as a single-disease management model in China (2025 Edition). Cancer Innov 2025; 4: e70031.

[4] Cai F, Luis MAF, Lin X, Wang M, Cai L, Cen C and Biskup E. Anthracycline-induced cardiotoxicity in the chemotherapy treatment of breast cancer: preventive strategies and treatment. Mol Clin Oncol 2019; 11: 15-23.

[5] Seoane JA, Kirkland JG, Caswell-Jin JL, Crabtree GR and Curtis C. Chromatin regulators mediate anthracycline sensitivity in breast cancer. Nat Med 2019; 25: 1721-1727.

[6] Xing M, Wang J, Yang Q, Wang Y, Li J, Xiong J and Zhou S. FKBP12 is a predictive biomarker for efficacy of anthracycline-based chemotherapy in breast cancer. Cancer Chemother Pharmacol 2019; 84: 861-872.
[7] Cassidy T, Humphries AR, Craig M and Mackey MC. Characterizing chemotherapy-induced neutropenia and monocytopenia through mathematical modelling. Bull Math Biol 2020; 82: 104.
[8] Ba Y, Shi Y, Jiang W, Feng J, Cheng Y, Xiao L, Zhang Q, Qiu W, Xu B, Xu R, Shen B, Luo Z, Xie X, Chang J, Wang M, Li Y, Shuang Y, Niu Z, Liu B, Zhang J, Zhang L, Yao H, Xie C, Huang H, Liao W, Chen G, Zhang X, An H, Deng Y, Gong P, Xiong J, Yao Q, An X, Chen C, Shi Y, Wang J, Wang X, Wang Z, Xing P, Yang S and Zhou C. Current management of chemotherapy-induced neutropenia in adults: key points and new challenges: Committee of Neoplastic Supportive-Care (CONS), China Anti-Cancer Association Committee of Clinical Chemotherapy, China Anti-Cancer Association. Cancer Biol Med 2020; 17: 896-909.
[9] Yin F, Xi YL, Wang Y, Li BR, Qian J, Ren H, Zhang J, Tang HZ and Ning BT. The clinical outcomes and biomarker features of severe sepsis/septic shock with severe neutropenia: a retrospective cohort study. Transl Pediatr 2021; 10: 464-473.
[10] Khouj E, Marafi D, Aljamal B, Hajiya A, Elshafie RM, Hashem MO, Abdulwahab F, Jaafar A, Alshidi T, Aboelanine AH, Awaji A and Alkuraya FS. Human 'knockouts' of CSF3 display severe congenital neutropenia. Br J Haematol 2023; 203: 477-480.
[11] Foley AM and Hoffman M. CE: febrile neutropenia in the chemotherapy patient. Am J Nurs 2023; 123: 36-42.
[12] Chen Y, Zhou H and Jiang J. Knowledge, attitude and practice towards chemotherapy-related neutropenia and febrile neutropenia among breast cancer patients. BMJ Open 2024; 14: e089738.
[13] Theyab A, Alsharif KF, Alzahrani KJ, Oyouni AAA, Hawsawi YM, Algahtani M, Alghamdi S and Alshammary AF. New insight into strategies used to develop long-acting G-CSF biologics for neutropenia therapy. Front Oncol 2023; 12: 1026377.
[14] Muzzana M, Pedrazzoli P and Lasagna A. G-CSF and G-CSF-related vasculitis: a systematic review of the literature and intriguing future research perspectives. Future Oncol 2021; 17: 4619-4634.
[15] Rahayatri TH, Oswari H, Kekalih A, Harahap A, Hendarto A, Munasir Z, Setiabudy R and Taher A. Granulocyte colony-stimulating factor improves innate immunity in pediatric pretransplant patients. J Clin Exp Hepatol 2024; 14: 101282.
[16] Zou W and Xu NL. Development and internal validation of a nomogram used to predict chemotherapy-induced neutropenia in non-small cell lung cancer patients: a retrospective cohort study. Cancer Manag Res 2021; 13: 2797-2804.
[17] Zhu Y, Guo D, Kong X, Liu S and Yu C. A risk-prediction nomogram for neutropenia or febrile neutropenia after etoposide-based chemotherapy in cancer patients: a retrospective cohort study. Pharmacology 2022; 107: 69-80.
[18] Tian T, Hu W and Hao J. Nomogram for predicting neutropenia in patients with esophageal, gastric, or colorectal cancer treated by chemotherapy in the first cycle. Int J Biol Markers 2024; 39: 23-30.
[19] Facon T, Kumar S, Plesner T, Orlowski RZ, Moreau P, Bahlis N, Basu S, Nahi H, Hulin C, Quach H, Goldschmidt H, O'Dwyer M, Perrot A, Venner CP, Weisel K, Mace JR, Raje N, Attal M, Tiab M, Macro M, Frenzel L, Leleu X, Ahmadi T, Chiu C, Wang J, Van Rampelbergh R, Uhlar CM, Kobos R, Qi M and Usmani SZ; MAIA Trial Investigators. Daratumumab plus lenalidomide and dexamethasone for untreated myeloma. N Engl J Med 2019; 380: 2104-2115.
[20] Peduzzi P, Concato J, Kemper E, Holford TR and Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996; 49: 1373-1379.
[21] Zhang H, Jia H, Wang H, Dong Q, Zhang Y and Hou Z. Development and validation of a nomogram for predicting fasciotomy requirement in lower extremity arterial injuries: a retrospective case-control study. J Orthop Traumatol 2025; 26: 73.
[22] Huang J, Lu S, Wang J, Jiang L, Luo X, He X, Wu Y, Wang Y, Zhu X, Chen J, Tang Y, Chen K, Tian X, Shi B, Guo L, Zhu J, Sun F, Zhen Z and Zhang Y. A multicenter phase II trial of primary prophylactic PEG-rhG-CSF in pediatric patients with solid tumors and non-Hodgkin lymphoma after chemotherapy: an interim analysis. Cancer Med 2023; 12: 14130-14137.
[23] Todaka A, Sasaki M, Ueno H, Goto T, Murohisa G, Mizuno N, Ozaka M, Kobayashi S, Uesugi K, Kobayashi N, Hayashi H, Sudo K, Okano N, Horita Y, Kamei K, Nanami S and Boku N. FOLFIRINOX in pancreatic cancer: risk factors for febrile neutropenia and severe neutropenia - nationwide study analysis. Anticancer Res 2023; 43: 4115-4123.
[24] Dale DC, Bolyard AA and Makaryan V. The promise of novel treatments for severe chronic neutropenia. Expert Rev Hematol 2023; 16: 1025-1033.

[25] Abdel-Rahman O. Effect of body mass index on 5-FU-based chemotherapy toxicity and efficacy among patients with metastatic colorectal cancer; a pooled analysis of 5 randomized trials. Clin Colorectal Cancer 2019; 18: e385-e393.
[26] Yang X, Ni H, Lu Z, Zhang J, Zhang Q, Ning S, Qi L and Xiang B. Mesenchymal circulating tumor cells and Ki67: their mutual correlation and prognostic implications in hepatocellular carcinoma. BMC Cancer 2023; 23: 10.
[27] Danzinger S, Spornberger VH, Vietzen H, Tendl-Schulz K, Pfeiler G, Singer CF and Seifert M. Influence of histopathological changes after neoadjuvant chemotherapy on the survival of breast cancer patients. Cancer Treat Res Commun 2025; 43: 100886.
[28] Krupa K, Liszcz-Tymoszuk A, Czerw N, Czerw A, Sygit K, Kozłowski R, Deptała A and Badowska-Kozakiewicz A. CDK4/6 as a therapeutic target in HR+/HER2- breast cancer cells-current treatment status. Cancers (Basel) 2025; 17: 1039.
[29] Sethakorn N, Heninger E, Sánchez-de-Diego C, Ding AB, Yada RC, Kerr SC, Kosoff D, Beebe DJ and Lang JM. Advancing treatment of bone metastases through novel translational approaches targeting the bone microenvironment. Cancers (Basel) 2022; 14: 757.
[30] Sun X, Li K, Aryal UK, Li B and Yokota H. PI3K-activated MSC proteomes inhibit mammary tumors via Hsp90ab1 and Myh9. Mol Ther Oncolytics 2022; 26: 360-371.