Discover Oncology

Analysis

A nomogram for predicting overall survival in young triple-negative breast cancer patients: a population-based study

Guangwen Zhang[1] · Xinle Wang[1] · Chen Cheng[1] · Shiming Wang[2] · Yujun Guo[2,3]

Received: 21 November 2024 / Accepted: 16 May 2025
Published online: 23 May 2025

Abstract

Background Young triple negative breast cancer (YTNBC) patients often face poor survival outcomes. Given the high-risk nature of YTNBC, there is an urgent need for tools that can accurately predict patient outcomes and guide personalized treatment strategies. Prognostic models, particularly those in the form of nomograms, have gained popularity in oncology for their ability to integrate multiple clinical variables to estimate individual patient survival. Our study aimed to investigate independent prognostic factors in YTNBC patients and develop a nomogram to predict OS, thereby helping patients choose a better therapeutic approach.

Methods Patients diagnosed with YTNBC between January 2010 and December 2015 from the Surveillance, Epidemiology, and End Results (SEER) database were enrolled and randomly divided into training and validation cohorts at a ratio of 7:3. Univariate and multivariate Cox analyses were conducted to identify significant factors associated with prognosis, which were then used to construct a nomogram for predicting 1-, 3-, and 5-year OS.

Results Nine survival predictors (marital status, tumor grade, AJCC stage, T stage, N stage, M stage, surgery, bone metastases, brain metastases) were selected for nomogram construction. The concordance indexes (C-index), in the training and validation cohorts were 0.749 and 0.745, respectively. The nomogram model demonstrated good calibration, and time-dependent receiver operating characteristic (ROC) curves confirmed its superiority for clinical utility. Additionally, Kaplan–Meier survival curves of various independent prognostic factors validated the model.

Conclusions The novel nomogram serves as a reliable tool for predicting survival, aiding clinicians in identifying high-risk patients and devising individualized treatments.

Keywords Young triple negative breast cancer · Nomogram · Prognosis · Overall survival · Prediction

1 Introduction

Breast cancer (BC) has become the second most prevalent cancer type globally, following lung cancer, and ranks as the fourth leading cause of cancer-related mortality as of 2022 [1]. Among the various subtypes of BC, triple negative breast cancer (TNBC) constitutes approximately 15–20% of all invasive BC cases [2]. TNBC is characterized by its aggressive biological and clinical behavior, featuring highly proliferative tumor cells and limited treatment options. It is recognized as a high-grade cancer with a dismal prognosis compared to other BC types [3, 4].

✉ Shiming Wang, wangshiming@sxmu.edu.cn; ✉ Yujun Guo, guoyujun@sydyy.com | [1]First Clinical Medical College of Shanxi Medical University, Taiyuan 030001, Shanxi, China. [2]Department of Breast Surgery, First Hospital of Shanxi Medical University, Taiyuan 030001, Shanxi, China. [3]The Paper's Lead Contact, Taiyuan, China.

Age serves as an independent risk factor for BC survival, with young breast cancer (YBC) patients typically experiencing poorer outcomes compared to older age groups [5–7]. The fifth edition of the European Society for Medical Oncology (ESMO) guidelines defines YBC patients as those under 40 years old [8]. According to data from the Surveillance, Epidemiology, and End Results (SEER) program, approximately 5.6% of all invasive BC cases occur in YBC patients, and this incidence is steadily increasing over time [9]. Notably, numerous studies have highlighted that young triple negative breast cancer (YTNBC) exhibits heightened aggressiveness, characterized by features such as high tumor grade and a higher frequency of Breast Cancer gene 1/2 (BRCA1/2) mutations [7, 10, 11]. This aggressive phenotype of YTNBC is strongly associated with a poorer prognosis, underscoring the urgent need for targeted and effective therapeutic interventions.

Accurately predicting prognosis in YTNBC is paramount for developing tailored treatment strategies. However, the traditional American Joint Committee on Cancer (AJCC) staging system [12] may not adequately predict survival probability in YTNBC patients. This is because it primarily considers histologic metastasis of the tumor and fails to incorporate numerous other crucial prognostic factors, including age, race, marital status, tumor size, and treatment information. Among these, treatment methods are crucial in predicting overall survival, as they not only directly impact patient outcomes and recovery but also have the potential to significantly alter the trajectory of individual prognosis. Xu et al. has indicated that breast surgery can significantly reduce the risk of death in TNBC patients [13]. Moreover, TNBC demonstrates heightened sensitivity to chemotherapy compared to other breast cancer subtypes, making chemotherapy a key therapeutic approach. Pathological complete response to chemotherapy strongly correlates with long-term survival outcomes [14]. However, these factors are seldom explored specifically in the context of YTNBC.

Hence, there is a pressing need for a new prognostic tool tailored to YTNBC patients that comprehensively considers the unique prognostic factors of YTNBC, to accurately assess prognosis and guide treatment decisions. In the current study, we aimed to explore the independent prognostic factors for YTNBC, for the purpose of determining the function of the controversial factors in YTNBC and developing a prognostic nomogram for YTNBC.

2 Materials and methods

2.1 Patients selection and study design

The Surveillance, Epidemiology, and End Results (SEER) database from the National Cancer Institute is a comprehensive population-based cancer registry covering approximately 30% of the United States population. For this study, data were extracted from the SEER 17 registry database using SEER*Stat 8.4.2 software. We specifically focused on young patients diagnosed with TNBC between 2010 and 2015. Our inclusion criteria were as follows: (1) patients diagnosed with invasive breast cancer, (2) females under the age of 40 years, and (3) breast cancer confirmed as the first primary tumor through histology. Exclusion criteria included: (1) diagnosis with inflammatory breast cancer or Paget's disease, (2) cases lacking follow-up records (indicated by a survival time code of 0 months), and (3) missing information on key variables including race, marital status, laterality, tumor grade, TNM stage, surgery type, and presence of bone, brain, liver, and lung metastasis. Ultimately, a total of 2,360 eligible patients were included in our study. To ensure robustness and reliability, we randomly divided these patients into a training cohort (n = 1,652) and a validation cohort (n = 708) at a ratio of 7:3, as illustrated in Fig. 1 [15–17]. This division ratio was chosen based on previous research practices, which have found the 7:3 ratio to be suitable for constructing and validating predictive models [15]. By allocating most of the data to the training cohort, we aimed to maximize the sensitivity and specificity of the nomogram construction, while reserving a smaller portion for validation.

2.2 Variable collection

The present study encompassed several variables: baseline demographics (including age at diagnosis, race, and marital status), tumor features (including tumor site, laterality, histological type, tumor grade, T stage, N stage, M stage, and AJCC stage), therapy information (comprising surgery, radiation, and chemotherapy), and survival variables (namely vital status and survival months). We reclassified all enrolled patients according to the eighth edition of the AJCC staging system [12, 18]. The selection of the age cutoff value was grounded on a previously published study [8]. In our investigation, the primary outcome of interest was overall survival (OS), delineated as the time span from the date of diagnosis to the date of death from any cause.

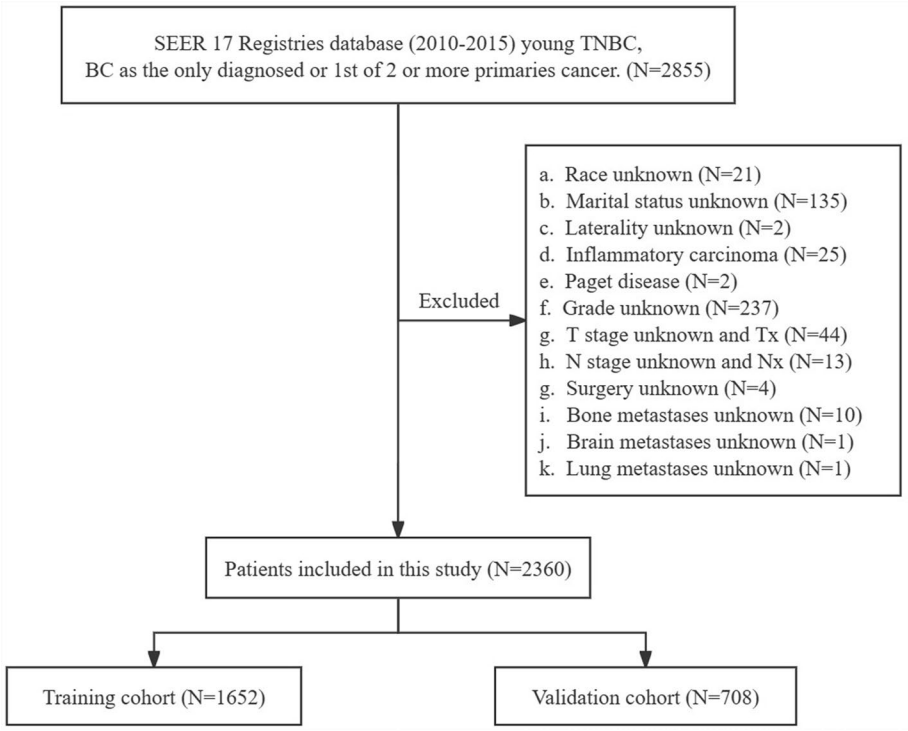

Fig. 1 The flowchart illustrates the process of patient selection in this study

2.3 Statistical analysis

The characteristics of both the training and validation cohorts were compared utilizing the chi-squared test. Univariate and multivariate Cox analyses were employed to identify independent prognostic factors. Subsequently, all identified independent risk factors were incorporated into the construction of nomograms for estimating 1-, 3-, and 5-year OS. The discriminative ability of the nomograms was assessed through the concordance index (C-index) and receiver operating characteristic (ROC) curves. Calibration plots were utilized to evaluate the predictive performance of the nomograms, estimating the concordance between predicted and observed survival probabilities. Additionally, Kaplan–Meier survival analyses were conducted to illustrate the survival duration associated with different risk factors.

In this study, SPSS 26.0 and R software (version 4.3.3) were adopted for all statistical analyses. All *P*-values were two-sided, and $P < 0.05$ was considered statistically significant.

3 Results

3.1 Patient characteristics

A total of 2,360 YTNBC participants were enrolled in this study, as outlined in Fig. 1, illustrating the specific screening process. Ultimately, 1,652 patients were assigned to the training cohort, while 708 patients comprised the validation cohort. The study population consisted of individuals under 40 years old, with 1,051 (44.5%) patients aged younger than 35. The median follow-up duration was 80 months, during which 559 (23.7%) patients had deceased. With a median follow-up time of 80 months, the 1-year, 3-year, and 5-year overall survival (OS) rates were 96.5%, 82.2%, and 77.6%, respectively. In the training cohort, over half of the patients were aged over 35 (56.3%), identified as white (69.6%), married (59.1%), and had tumors located on the left side (51.0%). Invasive ductal carcinoma (IDC) was the predominant pathological subtype of YTNBC, with 90.9% and 92.9% of cases classified as poorly differentiated (Grade III and IV) in the training and validation cohorts, respectively. Furthermore, the proportion of patients receiving radiation therapy was considerably lower compared to those undergoing surgery and chemotherapy, accounting for 46.2%, 92.3%, and 91.8% in the training cohort, respectively. Additionally, among TNBC patients, bone metastasis exhibited the highest incidence (1.9%), followed by lung metastasis (1.7%). Detailed demographic and clinicopathological characteristics of both the training and validation cohorts are presented in Table 1.

Importantly, there were no statistically significant differences observed in the distribution of all variables between the training and validation cohorts ($P > 0.05$).

3.2 Univariate and multivariate cox regression analysis

The results of the univariate Cox analysis are presented in Table 2. Fifteen variables, including race, marital status, tumor site, laterality, tumor grade, AJCC stage, T stage, N stage, M stage, surgery, radiation, bone metastasis, brain metastasis, liver metastasis, and lung metastasis, were found to be statistically associated with the OS of YTNBC patients. Subsequently, these fifteen variables were included in the multivariate analysis. The results of the multivariate analysis identified nine variables as final prognostic factors for OS (Table 2). These final prognostic factors include marital status, tumor grade, AJCC stage, T stage, N stage, M stage, surgery, bone metastasis, and brain metastasis, with statistical significance ($P < 0.05$).

3.3 Construction and validation of nomogram

The prognostic nomogram integrating the nine variables selected by multivariate Cox regression is depicted in Fig. 2. Notably, brain metastasis emerged as the most influential factor contributing to prognosis, followed by M stage, bone metastasis, N stage, AJCC stage, tumor grade, T stage, surgery, and marital status. Each subtype within these variables was assigned a score on the point scale. By summing up the total score and locating it on the total point scale, clinicians can easily draw a straight line down to estimate the probabilities of 1-year, 3-year, and 5-year OS.

The predictive accuracy of our model at 1-year, 3-year, and 5-year intervals, measured by the area under the curve (AUC) values in the internal validation, were 0.887, 0.808, and 0.767, respectively, as depicted by the receiver operating characteristic (ROC) curves (Fig. 3A). Notably, the 1-year, 3-year, and 5-year OS prediction curves of the nomogram closely approximated the ideal curves, where predicted OS matches observed OS. The Harrell's concordance index (C-index) was 0.749 (95% CI: 0.724 to 0.774) in the training cohort, indicating good discriminative ability. Furthermore, the calibration curves of the nomogram demonstrated high consistency between predicted and observed probabilities of OS in the training cohort (Fig. 4A–C), thereby validating the reliability of our model. In the validation cohort, the AUC values for predicting 1-year, 3-year, and 5-year OS were 0.881, 0.792, and 0.760, respectively (Fig. 3B). The C-index was 0.745 (95% CI: 0.707 to 0.784), indicating robust predictive performance. Additionally, calibration curves illustrated good agreement between predicted and observed outcomes in the validation cohort (Fig. 4D–F).

Decision curve analysis (DCA) was conducted to assess the clinical utility of the nomogram. The DCA curves indicated that the nomogram provided superior predictions for 1-year, 3-year, and 5-year OS, offering greater net clinical benefits compared to alternative strategies, as evidenced in both the training and validation cohorts (Fig. 5A–F).

3.4 Survival probability of the nomogram based on prognostic and treatment-related factors

According to the nomogram, we further investigated the survival benefit of grade, AJCC stage, T stage, N stage, M stage, surgery, bone metastasis, brain metastasis, liver metastasis, and lung metastasis for YTNBC patients. Compared with lower AJCC stage, T stage, N stage, and M stage, YTNBC patients with higher stage have lower OS ($P < 0.0001$) (Fig. 6A–D). Meanwhile, higher pathological grade correlates with poorer survival outcomes ($P < 0.05$) (Fig. 6E). As shown in Fig. 6F, primary site surgery significantly prolonged the OS of YTNBC patients ($P < 0.0001$). However, chemotherapy and radiation did not improve the OS of YTNBC patients ($P = 0.18, 0.19$) (Fig. 6G–H). When YTNBC patients experience distant metastases such as lung metastasis, bone metastasis, liver metastasis, brain metastasis, etc., they lead to poor prognosis and lower OS ($P < 0.0001$) (Fig. 6I–L).

4 Discussion

It's widely recognized that various sociodemographic characteristics and clinical pathological features closely influence the clinical prognosis of YTNBC patients. For instance, research indicates a bleaker outlook for YBC patients, who exhibit a higher incidence of the TNBC clinical pathology type and an increased likelihood of distant metastasis [19]. In comparison to patients without distant metastasis, those with metastatic disease to organs such as the lungs, brain, liver, or bones experience notably poorer survival outcomes [20]. Within the cohort of BC patients with distant metastasis, the median OS of the TNBC subtype tends to be shorter when compared to hormone receptor-positive (HR +) and human

Table 1 Demographics and clinicopathologic characteristics of the training and validation cohort

Variables	Overall (N = 2360)	Training cohort (N = 1652)	Validation cohort (N = 708)	*P* value
Age				
< 35	1051(44.5%)	722 (43.7%)	329 (46.5%)	0.233
≥ 35	1309(55.5%)	930 (56.3%)	379 (53.5%)	
Race				
White	1648(69.9%)	1150 (69.6%)	498 (70.3%)	0.095
Black	466(19.7%)	341 (20.6%)	125 (17.7%)	
Other	246(10.4%)	161 (9.8%)	85 (12.0%)	
Marital status				
Single	769(32.6%)	537 (32.5%)	232 (32.8%)	0.477
Married	1382(58.5%)	976 (59.1%)	406 (57.3%)	
Other	209(8.9%)	139 (8.4%)	70 (9.9%)	
Site				
Outer	1101(46.6%)	765 (46.3%)	336 (47.5%)	0.720
Inner	474(20.1%)	329 (19.9%)	145 (20.5%)	
Other	785(33.3%)	558 (33.8%)	227 (32.1%)	
Laterality				
Left	1207(51.1%)	843 (51.0%)	364 (51.4%)	0.900
Right	1153(48.9%)	809 (49.0%)	344 (48.6%)	
Histology				
IDC	2115(89.6%)	1468 (88.9%)	647 (91.4%)	0.077
Other	245(10.4%)	184 (11.1%)	61 (8.6%)	
Grade				
I ~ II	200(8.5%)	150 (9.1%)	50 (7.1%)	0.125
III ~ IV	2160(91.5%)	1502 (90.9%)	658 (92.9%)	
AJCC stage				
I ~ II	1821(77.2%)	1269 (76.8%)	552 (78.0%)	0.578
III ~ IV	539(22.8%)	383 (23.2%)	156 (22.0%)	
T stage				
T1 ~ T2	1896(80.3%)	1332 (80.6%)	564 (79.7%)	0.627
T3 ~ T4	464(19.7%)	320 (19.4%)	144 (20.3%)	
N stage				
N0	1315(55.7%)	924 (55.9%)	391 (55.2%)	0.786
N1 ~ N3	1045(44.3%)	728 (44.1%)	317 (44.8%)	
M stage				
M0	2255(95.6%)	1580 (95.6%)	675 (95.3%)	0.828
M1	105(4.4%)	72 (4.4%)	33 (4.7%)	
Surgery				
No	181(7.7%)	127 (7.7%)	54 (7.6%)	1.000
Yes	2179(92.3%)	1525 (92.3%)	654 (92.4%)	
Chemotherapy				
No/Unknown	180(7.6%)	135 (8.2%)	45 (6.4%)	0.150
Yes	2180(92.4%)	1517 (91.8%)	663 (93.6%)	
Radiation				
No/Unknown	1269(53.8%)	889 (53.8%)	380 (53.7%)	0.986
Yes	1091(46.2%)	763 (46.2%)	328 (46.3%)	
Bone metastasis				
No	2316(98.1%)	1621 (98.1%)	695 (98.2%)	1.000
Yes	44(1.9%)	31 (1.9%)	13 (1.8%)	
Brain metastasis				
No	2353(99.7%)	1646 (99.6%)	707 (99.9%)	0.682

Table 1 (continued)

Variables	Overall (N = 2360)	Training cohort (N = 1652)	Validation cohort (N = 708)	*P* value
Yes	7(0.3%)	6 (0.4%)	1 (0.1%)	
Liver metastasis				
No	2329(98.7%)	1632 (98.8%)	697 (98.4%)	0.636
Yes	31(1.3%)	20 (1.2%)	11 (1.6%)	
Lung metastasis				
No	2319(98.3%)	1625 (98.4%)	694 (98.0%)	0.680
Yes	41(1.7%)	27 (1.6%)	14 (2.0%)	

For race, 'other' includes American Indian, AK Native, Asian, and Pacific Islander; For marital status, 'other' consists of divorced, separated, widowed, and domestic partner; For site, 'other' includes nipple, central portion, axillary tail, and overlapping lesion of breast; For histology, IDC means invasive ductal carcinoma, 'other' includes invasive lobular carcinoma, pleomorphic carcinoma, adenocarcinoma, adenoid cystic carcinoma, solid carcinoma, apocrine adenocarcinoma, atypical medullary carcinoma, and metaplastic carcinoma; For grade, Grade I and II mean well-differentiated, grade III and IV mean poorly differentiated; *P* means difference between the training and validation cohort

epidermal growth factor receptor 2-positive (HER2 +) subtypes [21]. Notably, brain metastases were identified as one of the most powerful markers of poor OS in YTNBC patients. Brain metastases are often diagnosed through advanced imaging techniques such as MRI or CT scans, and they may present with neurological symptoms, including headaches, seizures, or cognitive changes. The time to development of brain metastases from the initial diagnosis of breast cancer varies, but a rapid progression to brain involvement can significantly worsen prognosis. The high propensity of YTNBC to develop brain metastases may be due to the aggressive nature of the disease and its tendency to evade conventional therapies, highlighting the need for more focused research on early detection and novel therapeutic strategies for managing brain involvement. Given the intricate interplay of prognostic factors impacting the survival of YTNBC patients, accurately predicting their outcomes poses a significant challenge. Consequently, we have developed a prognostic nomogram designed specifically to forecast the OS of YTNBC patients.

In this study, we identified several factors as independent predictors of OS among YTNBC patients. In terms of the degree of influence on patient prognosis, these factors include brain metastasis, M stage, bone metastasis, N stage, AJCC stage, tumor grade, T stage, surgery, and marital status. Further bolstering our findings, Tzikas AK et al. reported that YTNBC patients tend to have lower pathological grade differentiation compared to their elderly counterparts and are at a higher risk of developing both brain and liver metastases [7]. This underscores the significance of tumor grade in assessing prognosis and emphasizes the need for targeted interventions to address metastatic spread in YBC patients. Moreover, our study aligns with existing literature highlighting the association between marital status, AJCC staging, T stage, N stage, M stage, and OS in YTNBC [22]. Notably, AJCC staging serves as an important prognostic indicator, with lower stages generally predicting better outcomes. At the same time, T, N, and M—components of TNM staging—also have independent prognostic value. For example, even if the overall stage is the same, the impact of a larger tumor size on survival may differ from that of lymph node involvement. Analyzing these factors individually provides a more detailed understanding of their respective contributions to prognosis. Additionally, factors such as pathological grade differentiation and the performance of primary site surgery are recognized as pivotal in prognostic assessment. Marital status is associated with patient prognosis, with married patients typically experiencing more favorable disease outcomes. This correlation may be due to several factors, including more comprehensive disease care, stronger emotional and psychological support, and more reliable financial resources. These insights collectively underscore the multifaceted nature of prognostication in YTNBC and emphasize the importance of integrating various clinical and pathological parameters to inform treatment strategies and optimize patient outcomes.

Based on the results of multivariate Cox analysis, a nomogram was constructed to predict the survival of YTNBC patients. The nomogram developed in this study accurately estimates the prognosis of YTNBC patients and contributes to their clinical management. To enhance understanding of nomogram utilization, let's consider an example of a YTNBC patient: a female, unmarried, with stage IV disease, AJCC stage III, T3, N2, M0, who underwent surgery. The estimated 1-year, 3-year, and 5-year OS probabilities for this patient are approximately 88%, 40%, and 29%, respectively. Generally, early-stage TNBC can be effectively controlled through surgical intervention and postoperative adjuvant chemotherapy. However, after distant metastasis occurs, the primary treatment approach typically shifts towards palliative therapies such as chemotherapy aimed at maintaining quality of life and alleviating symptoms.

Table 2 Univariable and multivariable cox regression analyses of overall survival in the training cohort

Variables	Univariate Analysis		Multivariate Analysis	
	HR (95% CI)	*P* Value	HR (95% CI)	*P* Value
Age				
< 35	Reference		–	–
≥ 35	1.10 (0.90–1.35)	*P* = 0.334	–	–
Race				
White	Reference		Reference	
Black	1.38 (1.10–1.73)	*P* = 0.005	1.04 (0.81–1.33)	*P* = 0.768
Other	0.90 (0.63–1.29)	*P* = 0.580	1.01 (0.70–1.45)	*P* = 0.973
Marital status				
Single	Reference		Reference	
Married	0.69 (0.56–0.85)	*P* < 0.001	0.79 (0.63–0.99)	*P* = 0.042
Other	0.90 (0.63–1.29)	*P* = 0.569	0.95 (0.66–1.37)	*P* = 0.771
Site				
Outer	Reference		Reference	
Inner	1.03 (0.78–1.35)	*P* = 0.859	1.13 (0.85–1.50)	*P* = 0.395
Other	1.38 (1.11–1.72)	*P* = 0.004	1.22 (0.98–1.53)	*P* = 0.081
Laterality				
Left	Reference		Reference	
Right	0.80 (0.66–0.97)	*P* = 0.026	0.88 (0.72–1.07)	*P* = 0.201
Histology				
IDC	Reference		–	–
Other	0.71 (0.50–1.02)	*P* = 0.062	–	–
Grade				
I ~ II	Reference		Reference	
III ~ IV	0.66 (0.49–0.89)	*P* = 0.006	0.56 (0.41–0.76)	*P* < 0.001
AJCC stage				
I ~ II	Reference		Reference	
III ~ IV	4.90 (4.02–5.97)	*P* < 0.001	1.81 (1.31–2.51)	*P* < 0.001
T stage				
T1 ~ T2	Reference		Reference	
T3 ~ T4	3.53 (2.88–4.32)	*P* < 0.001	1.54 (1.17–2.02)	*P* = 0.002
N stage				
N0	Reference		Reference	
N1 ~ N3	3.54 (2.86–4.39)	*P* < 0.001	2.00 (1.52–2.63)	*P* < 0.001
M stage				
M0	Reference		Reference	
M1	10.49 (7.91–13.90)	*P* < 0.001	2.68 (1.62–4.41)	*P* < 0.001
Surgery				
No	Reference		Reference	
Yes	0.30 (0.23–0.39)	*P* < 0.001	0.62 (0.43–0.88)	*P* = 0.007
Chemotherapy				
No/Unknown	Reference		–	–
Yes	1.00 (0.70–1.45)	*P* = 0.983	–	–
Radiation				
No/Unknown	Reference		Reference	
Yes	1.30 (1.06–1.58)	*P* = 0.010	1.15 (0.92–1.44)	*P* = 0.208
Bone metastasis				
No	Reference		Reference	
Yes	18.88 (12.66–28.15)	*P* < 0.001	2.32 (1.32–4.09)	*P* = 0.004

Table 2 (continued)

Variables	Univariate Analysis		Multivariate Analysis	
	HR (95% CI)	*P* Value	HR (95% CI)	*P* Value
Brain metastasis				
No	Reference		Reference	
Yes	26.75 (11.80–60.63)	*P* <0.001	3.27 (1.30–8.18)	*P* = 0.012
Liver metastasis				
No	Reference		Reference	
Yes	8.99 (5.44–14.87)	*P* <0.001	0.72 (0.39–1.33)	*P* =0.297
Lung metastasis				
No	Reference		Reference	
Yes	9.52 (6.27–14.45)	*P* <0.001	1.22 (0.71–2.10)	*P* = 0.468

HR means hazard ratio; CI means confidence interval; For race, 'other' includes American Indian, AK Native, Asian, and Pacific Islander; For marital status, 'other' consists of divorced, separated, widowed, and domestic partner; For site, 'other' includes nipple, central portion, axillary tail, and overlapping lesion of breast; For histology, IDC means invasive ductal carcinoma, 'other' includes invasive lobular carcinoma, pleomorphic carcinoma, adenocarcinoma, adenoid cystic carcinoma, solid carcinoma, apocrine adenocarcinoma, atypical medullary carcinoma, and metaplastic carcinoma; For grade, Grade I and II mean well-differentiated, grade III and IV mean poorly differentiated

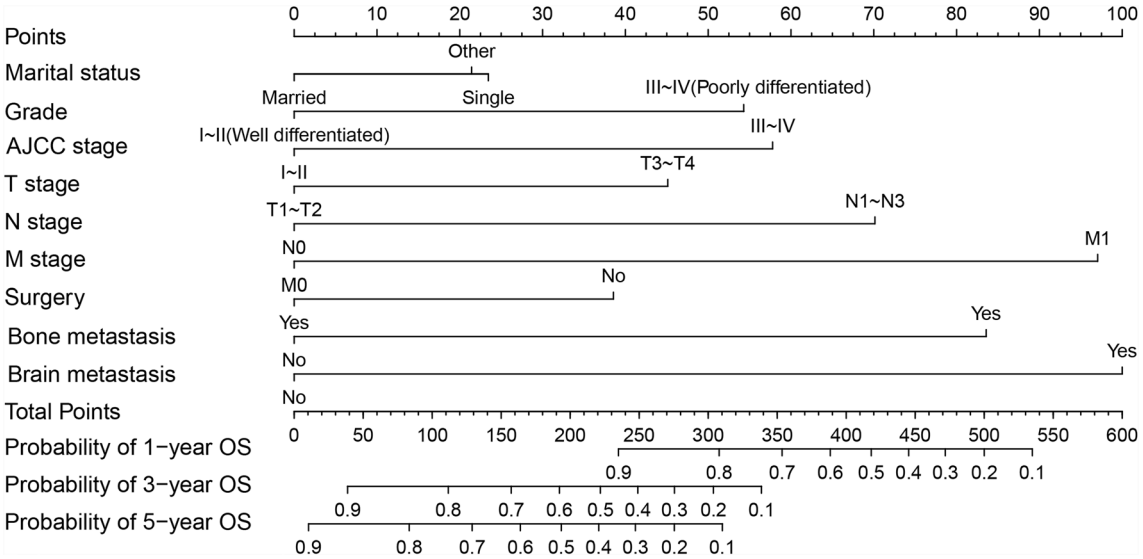

Fig. 2 Nomogram for predicting 1-, 3-, and 5-year overall survival (OS) for YTNBC patients

Accurately predicting the survival of YTNBC patients plays a pivotal role in guiding clinical practitioners and patients towards devising optimal treatment plans. This not only ensures the judicious utilization and allocation of medical resources but also enhances patient outcomes. When the predicted survival rate is favorable, clinicians may opt for more aggressive treatment strategies, aiming to maximize therapeutic efficacy. Conversely, if the predicted survival rate is poor, a palliative care approach supplemented by supportive therapies may be more prudent. This approach mitigates the potential side effects associated with aggressive treatments while prioritizing the improvement of the patient's quality of life. Furthermore, prognosticating the survival risk of YTNBC patients enables the customization of treatment plans, tailoring interventions to the individual patient's needs and circumstances. This personalized approach is instrumental in optimizing patient prognosis and fostering an improved quality of life. Thus, accurate survival prediction serves as a cornerstone in enhancing both clinical decision-making and patient-centered care in the management of YTNBC.

We employed a variety of methods to validate the clinical efficacy of the constructed nomogram. The predictive performance of the nomogram was evaluated based on its discrimination and internal calibration. The C-index of the nomogram demonstrates robust discriminatory ability, effectively distinguishing between different prognostic outcomes.

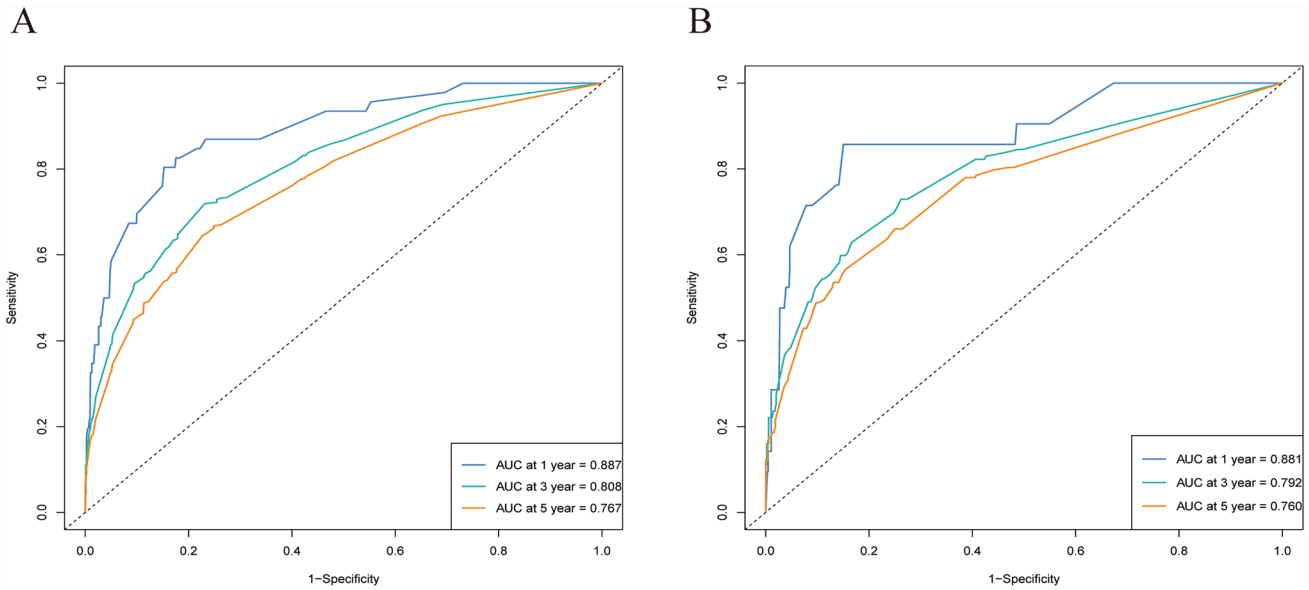

Fig. 3 Receiver operating characteristic (ROC) curves for survival prediction of YTNBC patients. **A** ROC curves of 1-, 3-, and 5-year OS in the training cohort; **B** ROC curves of 1-, 3-, and 5-year OS in the validation cohort. *AUC* the area under the time-dependent ROC curve

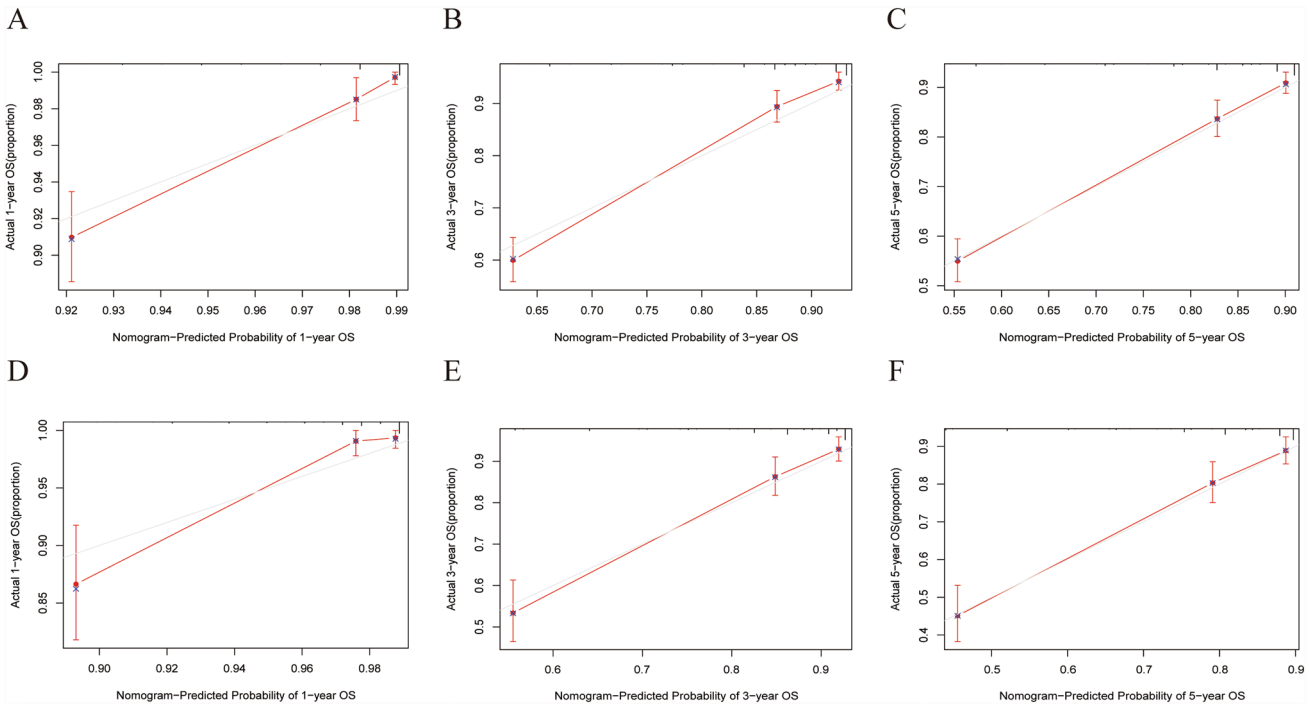

Fig. 4 Calibration curves for predicting 1-, 3-, and 5-year OS in the training cohort (**A**, **B**, **C**) and in the validation cohort (**D**, **E**, **F**)

Additionally, the AUC indicates that the nomogram has strong predictive capacity for YTNBC patient prognosis. The calibration curves exhibited excellent alignment between predicted and observed outcomes for OS, underscoring the reliability of the established nomogram. Moreover, to bolster the validation process, we conducted Kaplan–Meier curve analysis on independent prognostic factors for YTNBC patients. This comprehensive approach ensured the stability and accuracy of our predictive model, providing clinicians with a valuable tool for informing treatment decisions and optimizing patient care.

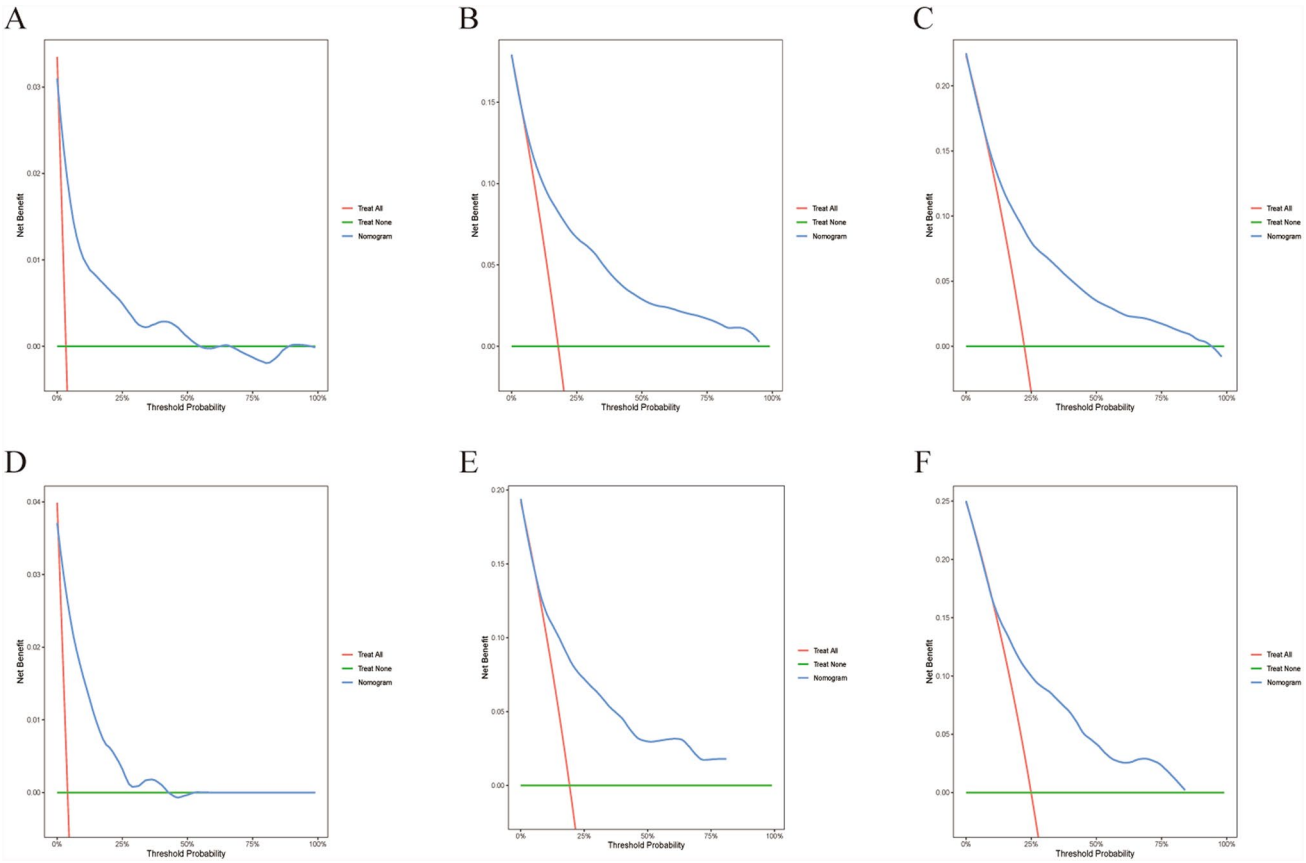

Fig. 5 Decision curve analyses (DCA) of the nomogram for 1-, 3-, and 5-year OS in the training cohort (**A**, **B**, **C**) and in the validation cohort (**D**, **E**, **F**)

In our study, primary site surgery significantly extends the OS of YTNBC patients. Currently, a comprehensive approach combining surgery, chemotherapy, radiotherapy, targeted therapy, and endocrine therapy proves beneficial for long-term survival in BC. For TNBC patients without distant metastasis, surgery remains a cornerstone of treatment, followed by adjuvant or neoadjuvant chemotherapy and radiotherapy to mitigate recurrence risks. Due to the negative expression of estrogen receptor (ER) and progesterone receptor (PR), as well as the negative or low expression of HER2 in TNBC, treatment options are limited because of the lack of clear targets. However, recent research has unveiled promising targets for TNBC, including immune and targeted therapy drugs such as anti-PD-1 antibodies, anti-PD-L1 antibodies, poly ADP-ribose polymerase (PARP) inhibitors, and antibody–drug conjugates (ADCs). The role of primary site surgery (mastectomy) in TNBC patients with distant metastasis remains contentious. Some studies suggest potential benefits for breast cancer patients with isolated bone metastases, while those with visceral metastases may not gain the same advantages [23, 24]. According to NCCN guidelines, surgery at the primary site is generally discouraged, except for patients who may benefit from initial systemic treatment. Thus, the treatment strategy for TNBC with distant metastasis typically involves a combination of chemotherapy, radiotherapy, immunotherapy, and targeted therapy.

The lack of improvement in OS following chemotherapy and radiation therapy observed in our study is more likely attributable to limitations inherent in the SEER database rather than a true lack of treatment efficacy. First, the SEER database does not provide critical information on treatment specifics such as chemotherapy regimens, dosages, administration timing, or radiotherapy target volumes. Without these details, it is difficult to fully assess and quantify the potential survival benefits of standard therapies. Moreover, in the SEER dataset, patients categorized as "No/Unknown" for chemotherapy or radiotherapy may include individuals who actually received treatment but lacked complete documentation, introducing misclassification bias and diluting the apparent benefits. Selection bias may also play a role, as patients with missing data or incomplete records—who were excluded from analysis—could include subgroups that are particularly responsive to chemoradiotherapy, further underestimating treatment impact. Importantly, numerous clinical studies have established the survival benefits of chemotherapy and radiotherapy in TNBC. For example, Lo et al. analyzed data from

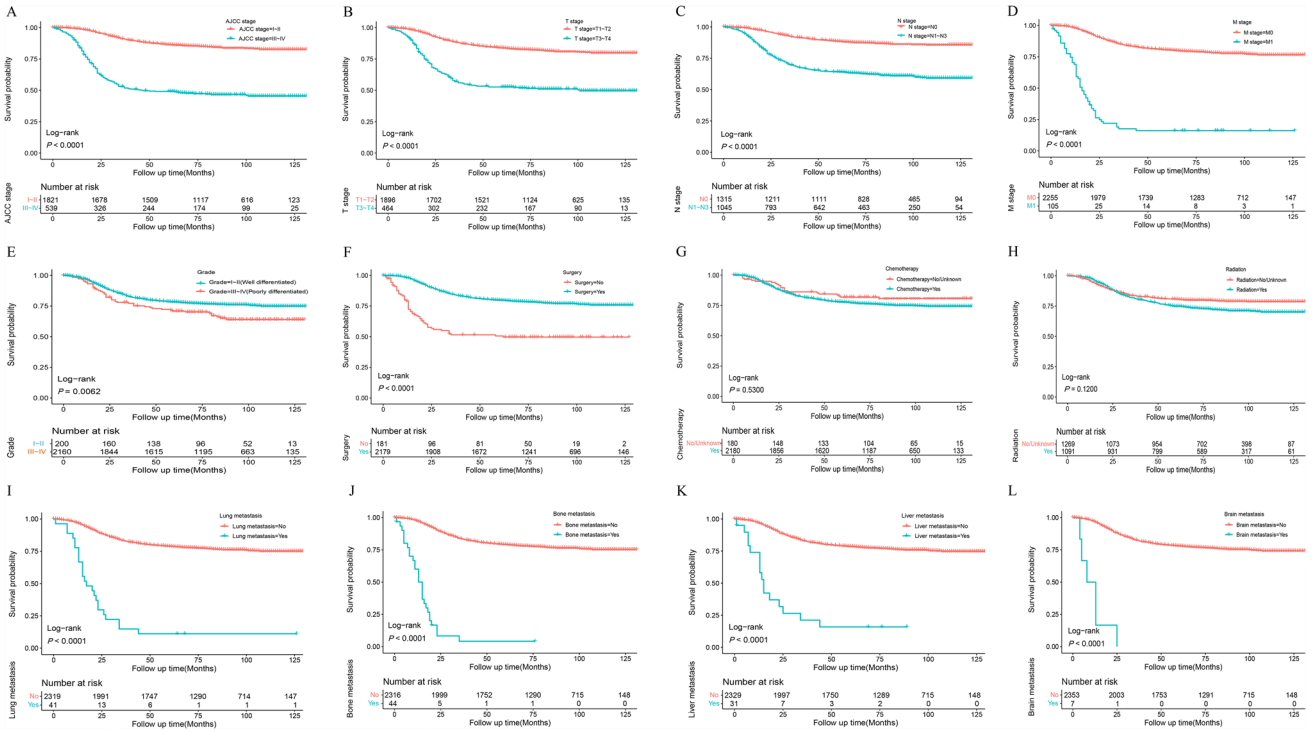

Fig. 6 Kaplan–Meier curves for various risk factors of YTNBC patients. AJCC stage (**A**), T stage (**B**), N stage (**C**), M stage (**D**), Grade (**E**), Surgery (**F**), Chemotherapy (**G**), Radiation (**H**), Lung metastasis (**I**), Bone metastasis (**J**), Liver metastasis (**K**), Brain metastasis (**L**)

the Taiwan Cancer Registry and found that adjuvant chemotherapy significantly improved both recurrence-free survival and overall survival in patients with T1cN0M0 TNBC [25]. Similarly, Xu et al. demonstrated that radiotherapy significantly improved overall survival and breast cancer-specific survival in elderly patients (≥ 70 years) with early-stage TNBC who underwent breast-conserving surgery, particularly among intermediate- and high-risk subgroups, based on SEER data and a validated prognostic nomogram [26]. These findings support current guidelines recommending chemotherapy and radiotherapy as essential components of TNBC treatment. Therefore, the lack of statistically significant OS benefit in our analysis should not be interpreted as a refutation of the effectiveness of these therapies, but rather as a reflection of the methodological and data-related constraints of large-scale population-based registries like SEER.

Undoubtedly, this study also suffers from several limitations that should be addressed in future research. First, the retrospective nature of the SEER database is a key limitation. While SEER provides a rich source of population-based data, the lack of randomized controlled trials or prospective follow-up introduces the possibility of bias in the selection of patients and treatments. Since the data are derived from clinical practice rather than experimental settings, there could be unmeasured confounders that affect the outcomes. Moreover, the SEER database only records information based on the time it was collected (2010–2015), so it may not accurately reflect the impact of newer therapies that have been introduced since then, such as immunotherapy and targeted therapies, which have significantly changed the landscape of treatment for YTNBC patients. For instance, Pembrolizumab, the first immune checkpoint inhibitor approved for TNBC, was approved in 2018. Additionally, Palbociclib, the first cyclin-dependent kinase 4 and 6 (CDK4/6) inhibitor approved for ER +, PR +, HER2- advanced or metastatic breast cancer, was approved in 2015. Secondly, there is inadequate information on lung, bone, liver, and brain metastases, including data on the type of distant metastatic lesions (solitary vs. multiple). Numerous retrospective studies have demonstrated that undergoing resection of solitary organ metastases can enhance long-term survival rates [27, 28]. The number of solitary organ metastases influences the decision for further surgical interventions, as solitary lesions in a single organ may be amenable to surgical resection. The absence of relevant information may compromise the accuracy of survival predictions by the model. Thirdly, other metastatic sites that could influence the prognosis of metastatic breast cancer, such as the peritoneum, other visceral organs, or skin, are not included in this study. Fourthly, due to limitations of the SEER database, factors such as patients'socioeconomic status, which often influence the likelihood of subsequent treatments, are not considered. Fifthly, given the large sample size included in this study, we employed the method of excluding missing values. However, while this approach

is straightforward, it may introduce bias, particularly when the missing data is not missing completely at random. To enhance the accuracy and reliability of the analysis, imputation techniques or model-based approaches are generally considered more appropriate strategies for handling missing data. Lastly, although our model demonstrates excellent predictive performance, it has yet to undergo external validation across diverse populations and in different centers or databases. This highlights the urgent need for further research to confirm its effectiveness and generalizability in a broader and more diverse patient cohort. External validation is crucial to ensure the model's consistent performance across various clinical settings and populations, thereby enhancing its applicability in real-world clinical practice. Due to the relatively low incidence of YTNBC, relying on individual centers for external validation is challenging. Therefore, we hope to establish databases for different populations in the future and conduct further validation to enhance the model's accuracy and applicability.

In conclusion, brain metastasis, bone metastasis, T stage, N stage, M stage, AJCC staging, surgery, tumor grade, and marital status stand out as independent factors impacting the prognosis of YTNBC. Developing a prognostic nomogram for YTNBC enables precise prediction of individual survival outcomes, empowering clinicians to tailor personalized treatment strategies for YTNBC patients. For example, incorporating patients'clinical information into the nomogram can predict their 1-year, 3-year, and 5-year survival rates, allowing for the early application of various treatment methods. This approach not only enhances the efficacy of therapeutic interventions but also fosters better patient outcomes and quality of life. Additionally, by integrating various clinical and pathological parameters, the nomogram facilitates a comprehensive understanding of YTNBC prognosis, paving the way for further advancements in precision medicine for this aggressive subtype of breast cancer.

5 Conclusions

We've identified crucial risk factors for YTNBC patients and integrated them into a nomogram for predicting patient prognosis. This prognostic model not only aids patients in gaining a clearer understanding of their future prognosis but also helps clinicians in tailoring personalized treatment strategies for individual YTNBC patients.

Acknowledgements We would like to thank SMW laboratory from Shanxi Medical University for their discussion. We appreciate the contributions of our colleagues who participated in the research and wrote this article.

Author contributions Guangwen Zhang: Writing–original draft, Supervision, Resources, Project administration, Formal analysis, Conceptualization. Xinle Wang: Software, Investigation, Data curation. Chen Cheng: Formal analysis, Data curation. Shiming Wang: Writing–review&editing, Resources, Investigation, Data curation. Yujun Guo: Writing–review&editing, Resources, Investigation, Data curation.

Funding This work was supported by the Natural Science Foundation of Shanxi Province [grant numbers 202203021221250].

Data availability The datasets generated and/or analyzed during the current study are available in the Surveillance, Epidemiology, and End Results (SEER) Program (https://seer.cancer.gov/). The data source is the SEER*Stat Database: Incidence – SEER Research Data, 17 Registries, Nov 2022 (2000–2020) – Linked to County Attributes – Time Dependent (1999–2021) Income/Rurality, 1969–2021 Counties, DCCPS, Surveillance Research Program, Released April 2023, based on the November 2022 submission.

Declarations

Ethics approval and consent to participate The data analyzed in this study were obtained from the SEER database, where patient identifiers were anonymized. Hence, this research was deemed exempt from Institutional Review Board approval.

Competing interests The authors declare no competing interests.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024. https://doi.org/10.3322/caac.21834.
2. Kumar P, Aggarwal R. An overview of triple-negative breast cancer. Arch Gynecol Obstet. 2016;293(2):247–69. https://doi.org/10.1007/s00404-015-3859-y.
3. Karim AM, Eun Kwon J, Ali T, et al. Triple-negative breast cancer: epidemiology, molecular mechanisms, and modern vaccine-based treatment strategies. Biochem Pharmacol. 2023;212: 115545. https://doi.org/10.1016/j.bcp.2023.115545.
4. Leon-Ferre RA, Goetz MP. Advances in systemic therapies for triple negative breast cancer. BMJ. 2023. https://doi.org/10.1136/bmj-2022-071674.
5. Walsh SM, Zabor EC, Flynn J, Stempel M, Morrow M, Gemignani ML. Breast cancer in young black women. Br J Surg. 2020;107(6):677–86. https://doi.org/10.1002/bjs.11401.
6. Gao X, Zhang F, Zhou Q, Xu H, Bian J. Metastasis, characteristic, and treatment of breast cancer in young women and older women: a study from the Surveillance, Epidemiology, and End Results registration database. PLoS ONE. 2023;18(11): e0293830. https://doi.org/10.1371/journal.pone.0293830.
7. Tzikas AK, Nemes S, Linderholm BK. A comparison between young and old patients with triple-negative breast cancer: biology, survival and metastatic patterns. Breast Cancer Res Treat. 2020;182(3):643–54. https://doi.org/10.1007/s10549-020-05727-x.
8. Paluch-Shimon S, Cardoso F, Partridge AH, et al. ESO-ESMO fifth international consensus guidelines for breast cancer in young women (BCY5). Ann Oncol. 2022;33(11):1097–118. https://doi.org/10.1016/j.annonc.2022.07.007.
9. Cathcart-Rake EJ, Ruddy KJ, Bleyer A, Johnson RH. Breast Cancer in adolescent and young adult women under the age of 40 years. JCO Oncol Pract. 2021;17(6):305–13. https://doi.org/10.1200/op.20.00793.
10. Vihervuori H, Korpinen K, Autere TA, Repo H, Talvinen K, Kronqvist P. Varying outcomes of triple-negative breast cancer in different age groups-prognostic value of clinical features and proliferation. Breast Cancer Res Treat. 2022;196(3):471–82. https://doi.org/10.1007/s10549-022-06767-1.
11. Ma D, Jiang YZ, Xiao Y, et al. Integrated molecular profiling of young and elderly patients with triple-negative breast cancer indicates different biological bases and clinical management strategies. Cancer. 2020;126(14):3209–18. https://doi.org/10.1002/cncr.32922.
12. Giuliano AE, Connolly JL, Edge SB, et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin. 2017. https://doi.org/10.3322/caac.21393.
13. Zhou HL, Chen DD. Prognosis of patients with triple-negative breast cancer: a population-based study from SEER database. Clin Breast Cancer. 2023;23(3):e85–94. https://doi.org/10.1016/j.clbc.2023.01.002.
14. Bianco N, Palazzo A, Pagan E, et al. Adjuvant treatment for triple negative breast cancer with residual tumor after neo-adjuvant chemotherapy a single institutional retrospective analysis. Breast. 2021;59:351–7. https://doi.org/10.1016/j.breast.2021.08.004.
15. Huang X, Luo Z, Liang W, et al. Survival nomogram for young breast cancer patients based on the SEER database and an external validation cohort. Ann Surg Oncol. 2022;29(9):5772–81. https://doi.org/10.1245/s10434-022-11911-8.
16. Wang W, Liu J, Chen Y, et al. An effective tool for predicting survival in breast cancer patients with de novo lung metastasis: nomograms constructed based on SEER. Front Surg. 2022;9: 939132. https://doi.org/10.3389/fsurg.2022.939132.
17. Li Y, Tao X, Ye Y, et al. Prognostic nomograms for young breast cancer: a retrospective study based on the SEER and METABRIC databases. Cancer Innov. 2024;3(6): e152. https://doi.org/10.1002/cai2.152.
18. 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. 2017. https://doi.org/10.3322/caac.21388.
19. Murphy BL, Day CN, Hoskin TL, Habermann EB, Boughey JC. Adolescents and young adults with breast cancer have more aggressive disease and treatment than patients in their forties. Ann Surg Oncol. 2019;26(12):3920–30. https://doi.org/10.1245/s10434-019-07653-9.
20. Guo Y, Arciero CA, Jiang R, Behera M, Peng L, Li X. Different breast cancer subtypes show different metastatic patterns: a study from a large public database. Asian Pac J Cancer Prev. 2020;21(12):3587–93. https://doi.org/10.31557/apjcp.2020.21.12.3587.
21. Criscitiello C, Marra A, Morganti S, et al. Clinical outcomes of patients with metastatic breast cancer enrolled in phase I clinical trials. Eur J Cancer. 2021;157:40–9. https://doi.org/10.1016/j.ejca.2021.07.037.
22. Zhang P, Xu BH, Ma F, Li Q. Clinicopathological characteristics and prognostic significance of young patients with triple-negative breast cancer. Zhonghua Zhong Liu Za Zhi. 2010;32(2):128–31.
23. Huang Z, Tan Q, Qin Q, Mo Q, Wei C. Impact of primary site surgery on survival of patients with de novo stage IV breast cancer. Cancer Manag Res. 2021;13:319–27. https://doi.org/10.2147/cmar.S280470.
24. Yu Y, Hong H, Wang Y, et al. Clinical evidence for locoregional surgery of the primary tumor in patients with De novo stage IV breast cancer. Ann Surg Oncol. 2021;28(9):5059–70. https://doi.org/10.1245/s10434-021-09650-3.
25. Lo C, Chang DY, Lu YS, et al. Benefit of adjuvant chemotherapy for T1cN0M0 and selected T1bN0M0 triple-negative breast cancer: a nationwide cancer registry-based study. Oncologist. 2025. https://doi.org/10.1093/oncolo/oyae346.
26. Xu Z, Qin C, Cao B, Ruan P, Zhang M, Chen G. The potential benefits of radiotherapy in elderly patients with early-stage triple-negative breast cancer. Front Med. 2024;11:1525425. https://doi.org/10.3389/fmed.2024.1525425.
27. Endoh M, Shiono S, Yamauchi Y, et al. Pulmonary metastasectomy for pulmonary metastasis of breast cancer has a limited prognostic impact: a multi-institutional retrospective analysis. J Thorac Dis. 2020;12(11):6552–62. https://doi.org/10.21037/jtd-20-1788.
28. He X, Zhang Q, Feng Y, et al. Resection of liver metastases from breast cancer: a multicentre analysis. Clin Transl Oncol. 2020;22(4):512–21. https://doi.org/10.1007/s12094-019-02155-2.