Cancer Pathogenesis and Therapy 1 (2023) 253–261

Contents lists available at ScienceDirect

Cancer Pathogenesis and Therapy

journal homepage: www.journals.elsevier.com/cancer-pathogenesis-and-therapy

Research article

Survival nomogram for patients with *de novo* metastatic breast cancer based on the SEER database and an external validation cohort

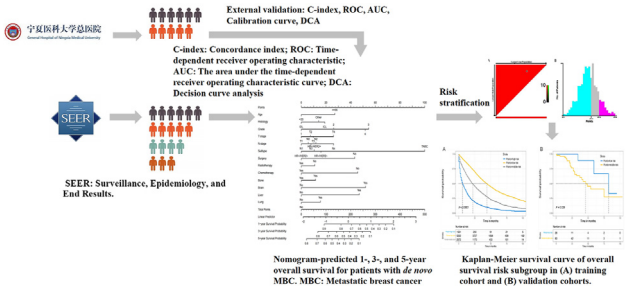
Lizhi Ning [a,1], Yaobang Liu [b,1], Yujin Hou [a], Miaozhou Wang [c], Mingqiang Shi [c], Zhen Liu [c], Jiuda Zhao [c,*], Xinlan Liu [a,**]

[a] *Department of Medical Oncology, General Hospital of Ningxia Medical University, Yinchuan, Ningxia 750004, China*
[b] *Department of Surgical Oncology, General Hospital of Ningxia Medical University, Yinchuan, Ningxia 750004, China*
[c] *Breast Disease Diagnosis and Treatment Center of Affiliated Hospital of Qinghai University & Affiliated Cancer Hospital of Qinghai University, Xining, Qinghai 810000, China*

HIGHLIGHTS

- A nomogram based on the Surveillance, Epidemiology, and End Results (SEER) database can predict 1-, 3-, and 5-year overall survival (OS) for patients with *de novo* metastatic breast cancer.
- The nomogram is universal by using external validation data from China.
- The nomogram is useful to predict OS and guide individualized therapy in patients with *de novo* metastatic breast cancer.

GRAPHICAL ABSTRACT



ARTICLE INFO

Managing Editor: Peng Lyu

Keywords:
De novo metastatic breast cancer
Overall survival
SEER
Nomogram

ABSTRACT

Background: On average, 5–10% of patients are diagnosed with metastatic breast cancer (MBC) at the initial diagnosis. This study aimed to develop a nomogram to predict the overall survival (OS) in these patients.

Methods: The nomogram was based on a retrospective study of 9435 patients with *de novo* MBC from the Surveillance, Epidemiology, and End Results (SEER) database. The predictive accuracy and discriminative ability of the nomogram were determined using the concordance index (C-index), area under the time-dependent receiver operating characteristic curve (AUC), and calibration curve. Decision curve analysis (DCA) was employed to evaluate the benefits and advantages of our new predicting model over the 8th edition of the American Joint Committee on Cancer (AJCC) Tumor Node Metastasis (TNM) staging system. The results were validated in a retrospective study of 103 patients with *de novo* MBC from January 2013 to June 2022 at an institution in northwest China.

Results: Multivariate analysis of the primary cohort revealed that independent factors for survival were age at diagnosis, pathological type, histological grade, T stage, N stage, molecular subtype, bone metastasis, brain metastasis, liver metastasis, lung metastasis, surgery, chemotherapy, and radiotherapy. The nomogram achieved a C-index of 0.688 (95% confidence interval [CI], 0.682–0.694) in the training cohort and 0.875 (95% CI, 0.816–0.934) in the validation cohort. The AUC of the nomograms indicated good specificity and sensitivity in the

* Corresponding author: Breast Disease Diagnosis and Treatment Center of Affiliated Hospital of Qinghai University & Affiliated Cancer Hospital of Qinghai University, Xining, Qinghai 810000, China.
** Corresponding author: Department of Medical Oncology, General Hospital of Ningxia Medical University, Yinchuan, Ningxia 750004, China.
E-mail addresses: jiudazhao@126.com (J. Zhao), nxliuxinlan@126.com (X. Liu).
[1] Lizhi Ning and Yaobang Liu contributed equally to this work.

https://doi.org/10.1016/j.cpt.2023.07.004
Received 12 April 2023; Received in revised form 12 July 2023; Accepted 23 July 2023

training and validation cohorts, respectively. Calibration curves showed favorable consistency between the predicted and actual survival probabilities. Additionally, the DCA curve produced higher net gains than by the AJCC-TNM staging system. Finally, risk stratification can accurately identify groups of patients with *de novo* MBC at different risk levels.

Conclusions: The nomogram showed favorable predictive and discriminative abilities for OS in patients with *de novo* MBC. Other populations from different countries or prospective studies are needed to further validate the nomogram.

Introduction

The Global Cancer Observatory (GLOBOCAN) estimates the incidence and mortality worldwide for 36 cancers in 185 countries. Female breast cancer has surpassed lung cancer as the most commonly diagnosed cancer, ranking first in the incidence rate of most countries and mortality rate of 110 countries.[1] On average, 5–10% of patients are diagnosed with stage IV disease at initial diagnosis, while 20–30% of patients with stage I–III eventually develop disease progression and distant metastasis.[2] Unlike early breast cancer, metastatic breast cancer (MBC) is incurable. Therefore, the main goals of treatment are symptomatic relief, improved quality of life, and increased survival.[3,4] It is generally accepted that systemic therapy, including chemotherapy, endocrine therapy, and targeted therapy, is the main and effective treatment for MBC.[5] Surgery for MBC is usually performed when the patient has complications such as skin ulcers, bleeding, and pain.

Recently, advances in local and systemic adjuvant therapy continue to reduce the risk of breast cancer recurrence and subsequent death.[6] However, due to the heterogeneity of MBC, it is especially important to accurately assess the prognosis of each patient for clinical treatment and guideline formulation.[7] At present, the most widely used staging in clinics is the American Joint Committee on Cancer (AJCC) Tumor Node Metastasis (TNM) staging system. However, the TNM classification system has been unable to meet the needs of clinicians due to the lack of important pathological indicators reflecting the prognosis of patients and clinical parameters, such as psychological and social factors. Nomograms incorporating the TNM classification system and other major clinical factors have been gradually used as visible and accurate models for survival prediction and risk stratification of various tumors.[8,9]

Nomograms are effective prognostic tools for predicting morbidity and survival. Nomograms quantitatively predict patient outcomes by integrating key prognostic factors and visualizing results in easy-to-understand graphs.[10] Several nomograms have been developed to predict the risk of developing relapsed MBC.[11–13] However, many studies have shown that the clinicopathological features of *de novo* MBC are distinct from those of patients with relapsed breast cancer, and patients with *de novo* MBC usually have better survival than those with recurrent/MBC.[14,15] The above differences may be explained by the fact that patients with recurrent/MBC often use adjuvant therapy for a longer duration and more frequent treatments during the course of the disease, resulting in greater resistance to treatment in recurrent MBC. Therefore, *de novo* MBC should not be confused with recurrent MBC. Zhao et al. constructed a nomogram for patients with *de novo* MBC based on the Surveillance, Epidemiology, and End Results (SEER) database; however, the validation cohort was also selected from the SEER database and may not be applicable to the Chinese population.[16] Our study constructed a nomogram containing multiple factors based on the SEER database and validated it using data from patients with *de novo* MBC in an institution in northwest China. The nomogram is universal, predicts prognosis, and guides individualized treatment for patients with *de novo* MBC.

Methods

Population selection

Data were obtained from the SEER database and an institution in northwest China. SEER is an open-access resource for tumor-based demographic and pathological information as well as treatment information and patient survival outcomes and consists of 18 population-based cancer registries. SEER*Stat Version 8.3.4 (http://www.seer.cancer.gov/seerstat) was used to identify eligible patients.

As the SEER database began collecting information on human epidermal growth factor receptor-2 (HER2) status and sites of distant metastasis in 2010, we collected data from patients with *de novo* MBC between 2010 and 2015. The inclusion criteria for patients with MBC were as follows: female, older than 18 years when diagnosed, breast cancer as the first and only malignant tumor diagnosis, histology of invasive ductal or lobular carcinoma (IDC, ILC), and at least one distant site of *de novo* metastasis. Patients with unknown differentiation grade, T stage, N stage, molecular subtype, site of metastasis, surgery, radiation, chemotherapy, or follow-up information were excluded. Patients in T0

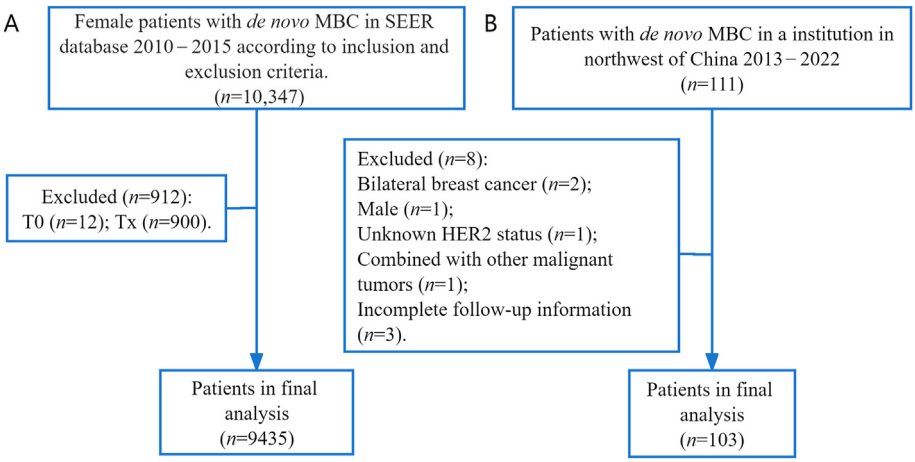


Figure 1. The flowchart of the patient selection process. (A) The training cohort included 9435 patients with *de novo* MBC. (B) The validation cohort included 103 patients with *de novo* MBC from an institution in northwest China. MBC: Metastatic breast cancer.

and Tx stages ($n = 912$) were also excluded to match the variables in the training and validation cohorts.

Based on the inclusion and exclusion criteria for the training cohort, an independent continuous cohort of patients from Ningxia Medical University General Hospital from January 2013 to June 2022 was included as the validation cohort.

Variable collection

We collected demographic features (including age at diagnosis), clinicopathological characteristics (including histological type, T stage, N stage, bone metastasis, lung metastasis, liver metastasis, brain metastasis, molecular subtype, and laterality), treatment (including breast surgery, radiotherapy, and chemotherapy), and survival data (including survival months and vital status) of each case. Patients were assigned into two groups according to the age of diagnosis: <55 years and ≥55 years. Patients were classified as having IDC (code: 8500/3), ILC (code: 8522/3), or others according to the International Classification of Diseases for Oncology third edition (ICD-O-3). Based on the expression of hormone receptor (HR) and HER2, patients were classified as having HR+/HER2-, HR+/HER2+, HR-/HER2+, and triple-negative breast cancer (TNBC). TNM stage classification was based on the AJCC and Breast Cancer system (8th edition). The main outcome of this study was overall survival (OS). OS was defined as the interval (months) from diagnosis to death from any cause, with loss of follow-up as censored data.

Statistical analysis

We performed a descriptive analysis of the demographic and clinicopathological features of the patients included in the training and validation cohorts. Parameters with a *P* value less than 0.1 in univariate Cox analysis or with clinical consideration of potential prognostic factors were included in the multivariable Cox model to identify independent prognostic factors in the training cohort. A nomogram to predict the 1-, 3-, and 5-year OS was constructed based on independent prognostic factors. The performance of the nomogram was evaluated using training and validation sets. The concordance index (C-index), time-dependent receiver operating characteristic (ROC) curve, and area under the ROC curve (AUC) were used to evaluate the distinguishing ability of the nomogram. The C-index and AUC values range from 0 to 1; higher values indicate a stronger predictive ability and values between 0.7 and 0.9 are generally considered to have good identification ability. Calibration curves were used to evaluate the accuracy of the point estimates of the nomogram-predicted survival with actual survival. The bootstrap resampling method (B = 1000) was used to plot the calibration curve. Decision curve analysis (DCA) was used to evaluate the advantages of our new prediction model over the 8th edition of the AJCC-TNM staging system. Survival curves were plotted using the Kaplan–Meier method. All analyses were performed using SPSS (version 26.0) and figure plots were generated using R software version 4.1.3 (www.r-project.org). X-tile software was used to calculate the optimal cutoff value. All statistical analyses were two-sided, and a *P* value of less than 0.05 was considered statistically significant.

Table 1
Baseline characteristics of patients with *de novo* metastatic breast cancer in the training and validation cohorts.

Characteristics	Training cohort ($n = 9435$)	Percent (%)	Validation cohort ($n = 103$)	Percent (%)
Age (years)				
<55	1494	15.8	44	42.7
≥55	7941	84.2	59	57.3
T stage				
T1	1402	14.9	11	10.7
T2	3152	33.4	40	38.8
T3	1596	16.9	9	8.7
T4	3285	34.8	43	41.7
N stage				
N0	2334	24.7	6	5.8
N1	4000	42.4	34	33.0
N2	1170	12.4	21	20.4
N3	1522	16.1	38	39.9
Nx	409	4.3	4	3.9
Laterality				
Left	4856	51.5	58	56.3
Right	4579	48.5	45	43.7
Grade				
I	829	8.8	8	7.8
II	4019	42.6	71	69.0
III	4520	47.9	12	11.7
IV	67	0.7	12	11.7
Histological type				
Invasive ductal carcinoma	7522	79.7	91	88.3
Invasive lobular carcinoma	993	10.5	4	3.9
Other	920	9.8	8	7.8
Chemotherapy				
Yes	5106	54.1	97	94.2
No	4329	45.9	6	5.8
Radiotherapy				
Yes	1823	19.3	26	25.2
No	7612	80.7	77	74.8
Surgery				
Yes	3629	38.5	41	39.8
No	5806	61.5	62	60.2
Molecular subtype				
HR+/HER2-	5778	61.2	48	46.6
HR+/HER2+	1455	15.4	22	21.4
HR-/HER2+	793	8.4	25	24.3
TNBC	1409	14.9	8	7.8
Bone				
Yes	6011	63.7	51	49.5
No	3424	36.3	52	50.5
Brain				
Yes	612	6.5	2	1.9
No	8823	93.5	101	98.1
Liver				
Yes	2177	23.1	26	25.2
No	7258	76.9	77	74.8
Lung				
Yes	3049	32.3	28	27.2
No	6386	67.7	75	72.8
Status				
Alive	2515	26.7	72	69.9
Dead	6920	73.3	31	30.1

HER2: Human epidermal growth factor receptor 2; HR: Hormone receptor; MBC: Metastatic breast cancer; TNBC: Triple-negative breast cancer.

Results

Patient characteristics

Ultimately, we identified 9435 patients from 2010 to 2015 in the SEER training cohort and 103 patients from an institution in northwest China as the validation cohort. The inclusion process is illustrated in Figure 1. The clinicopathological features and treatments of all patients are shown in Table 1. In the training cohort, the most common site of metastasis was the bone (63.7%), while the least common was the brain (6.5%). There were 23.1% and 32.3% of the patients with liver and lung metastases, respectively. HR+/HER2- was the most common subtype among patients with *de novo* MBC, making up 61.2%, while HR-/HER2+ was the least common (8.4%) subtype. HR+/HER2+ and TNBC subtypes comprised 15.4% and 14.9% of cases, respectively. Of these patients, 38.5% underwent surgery for the primary tumor. A total of 19.3% of patients received radiotherapy and 54.1% received chemotherapy. The median follow-up was 66.0 months (95% confidence interval [CI], 64.8–67.2 months). The 1-, 3-, and 5-year OS rates were 71.3%, 43.2%, and 27.3%, respectively.

In the validation cohort, 49.5%, 25.2%, 27.2%, and 1.9% of patients had bone, liver, lung, and brain metastases, respectively. The most

common subtype was HR+/HER2- (46.6%), whereas the least common subtype was TNBC (7.8%). A total of 39.8% of patients underwent surgery for the primary tumor, 25.2% received radiotherapy, and 94.2% received chemotherapy. The median follow-up time was 31 months (95% CI, 24.7–37.3 months) and the 1-, 3-, and 5-year OS rates were 97.0%, 68.5%, and 45.3%, respectively.

Nomogram construction

According to the univariate analysis of the training cohort, age at diagnosis, differentiation grade, T stage, N stage, molecular subtype, brain metastasis, liver metastasis, lung metastasis, surgery, radiotherapy, and chemotherapy were associated with OS ($P < 0.05$) [Table 2]. Considering the impact of bone metastasis on the prognosis of patients with breast cancer, bone metastasis was included in the Cox model. Of the factors incorporated into the multivariate Cox analysis, age at diagnosis, differentiation grade, T stage, N stage, subtype, bone metastasis, brain metastasis, liver metastasis, lung metastasis, surgery, radiotherapy, and chemotherapy were significantly associated with OS ($P < 0.05$) [Table 2]. These variables were incorporated into nomograms. Figure 2 shows the prediction of 1-, 3-, and 5-year OS probabilities in the nomogram. The specific value for each of these factors was assigned a score on a point scale and added to calculate the total score. Total points were used to estimate the 1-, 3-, and 5-year survival probabilities for each patient.

Nomogram validation and calibration

The nomogram was validated in the training and validation cohorts. The C-index was 0.688 (95% CI, 0.682–0.694) and 0.875 (95% CI,

Table 2
Univariate and multivariate Cox-regression analyses of factors associated with overall survival in the training cohort.

Variables	Univariate Cox			Multivariate Cox		
	HR	95% CI	P^a	HR	95% CI	P^a
Age (years)			<0.001[b]			<0.001[b]
<55		Reference			Reference	
≥55	1.373	1.283–1.470		1.339	1.250–1.435	
T stage			<0.001[b]			<0.001[b]
T1		Reference			Reference	
T2	1.086	1.006–1.172		1.084	1.004–1.171	
T3	1.164	1.068–1.269		1.111	1.017–1.213	
T4	1.478	1.372–1.593		1.326	1.226–1.434	
N stage			<0.001[b]			<0.001[b]
N0		Reference			Reference	
N1	0.972	0.916–1.032		0.942	0.886–1.002	
N2	0.957	0.882–1.038		1.070	0.982–1.165	
N3	0.949	0.880–1.024		1.025	0.947–1.109	
Nx	1.504	1.340–1.688		1.245	1.108–1.399	
Laterality			0.593			-
Left		Reference			Reference	
Right	0.982	0.917–1.051		-	-	
Grade			<0.001[b]			<0.001[b]
I		Reference			Reference	
II	1.286	1.171–1.412		1.328	1.208–1.461	
III	1.700	1.551–1.865		1.801	1.631–1.989	
IV	1.936	1.466–2.557		1.751	1.319–2.324	
Histological type			<0.001[b]			<0.001[b]
Invasive ductal carcinoma		Reference			Reference	
Invasive lobular carcinoma	0.965	0.894–1.042		1.228	1.132–1.333	
Other	1.239	1.146–1.339		1.161	1.073–1.257	
Chemotherapy			<0.001[b]			<0.001[b]
No		Reference			Reference	
Yes	0.724	0.691–0.759		0.615	0.583–0.649	
Radiotherapy			<0.001[b]			0.001[b]
No		Reference			Reference	
Yes	0.643	0.604–0.685		0.889	0.828–0.956	
Surgery			<0.001[b]			<0.001[b]
No		Reference			Reference	
Yes	0.587	0.559–0.618		0.628	0.591–0.666	
Molecular subtype			<0.001[b]			<0.001[b]
HR+/HER2-		Reference			Reference	
HR+/HER2+	0.837	0.780–0.898		0.839	0.778–0.906	
HR-/HER2+	0.931	0.850–1.021		0.938	0.849–1.037	
TNBC	2.205	2.069–2.350		2.452	2.277–2.640	
Bone			0.127			<0.001[b]
No		Reference			Reference	
Yes	1.039	0.989–1.092		1.131	1.073–1.193	
Brain	-	-	<0.001[b]			<0.001[b]
No		Reference			Reference	
Yes	2.084	1.909–2.275		1.741	1.591–1.904	
Liver			<0.001[b]			<0.001[b]
No		Reference			Reference	
Yes	1.578	1.495–1.666		1.653	1.562–1.750	
Lung			<0.001[b]			<0.001[b]
No		Reference			Reference	
Yes	1.380	1.313–1.450		1.179	1.119–1.243	

[a] P value was calculated by the Cox-regression test.
[b] $P < 0.05$. HER2: Human epidermal growth factor receptor 2; HR: Hormone receptor; OS: Overall survival; TNBC: Triple-negative breast cancer.

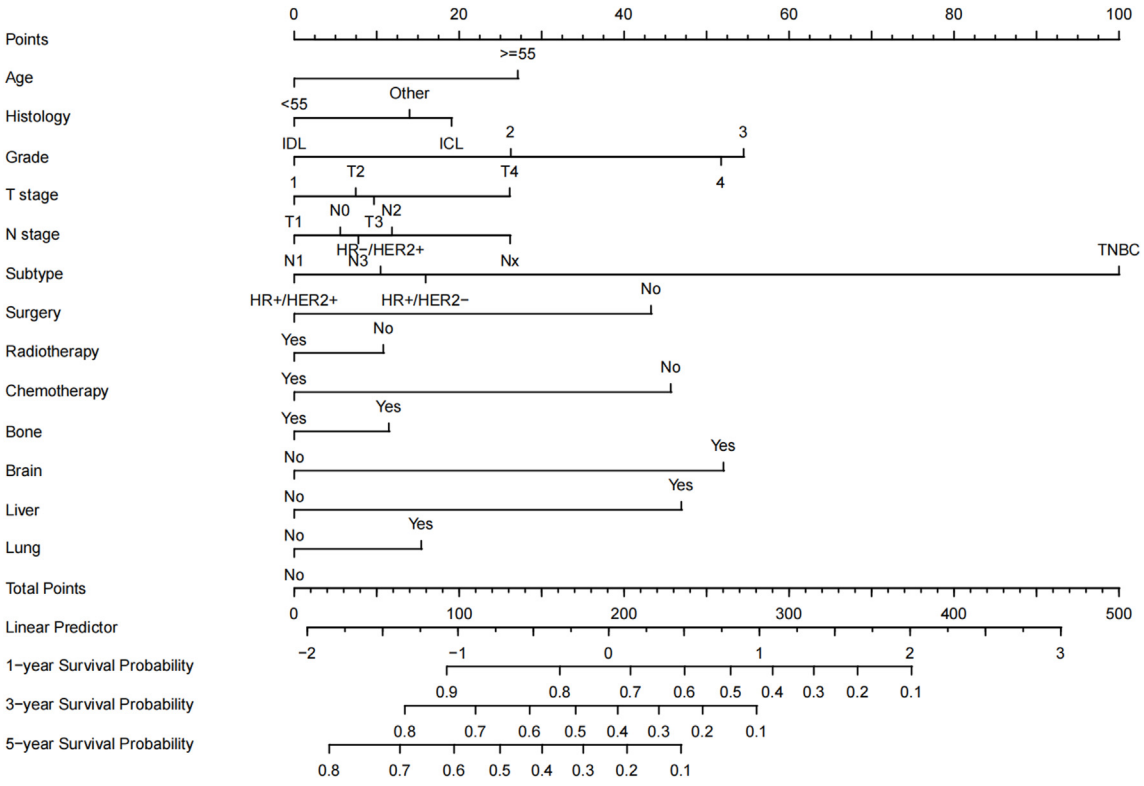


Figure 2. Nomogram-predicted 1-, 3-, and 5-year overall survival for patients with *de novo* MBC. MBC: Metastatic breast cancer.

0.816–0.934) in the training and validation cohort, respectively. In the training cohort, the AUC values of the nomogram for predicting 1-, 3-, and 5-year OS were 0.760 (95% CI, 0.740–0.780), 0.700 (95% CI, 0.670–0.730), and 0.690 (95% CI, 0.660–0.740), respectively. In the validation cohort, the AUC values of the nomogram for predicting 1-, 3-, and 5-year OS were 0.910 (95% CI, 0.880–0.920), 0.790 (95% CI, 0.770–0.850), and 0.860 (95% CI, 0.830–0.900), respectively [Figure 3]. The 1-, 3-, and 5-year calibration curves of the nomogram for OS prediction demonstrated good consistency between the training and validation cohorts [Figure 4]. As shown in Figure 5, the DCA indicated the growth of the net benefits of the new model over the 8th version of the AJCC-TNM staging system, with wide and practical ranges of threshold probabilities.

Risk stratification of patients with de novo metastatic breast cancer

The total OS score corresponding to each patient in the modeling set was obtained according to the established nomogram, and the optimal cutoff value related to prognosis was calculated using X-tile software. The patients were artificially divided into low-risk (score ≤257.5), medium-risk (257.5 < score ≤321.9), and high-risk groups (score >321.9) by the OS nomogram score [Figure 6]. We then plotted the Kaplan–Meier

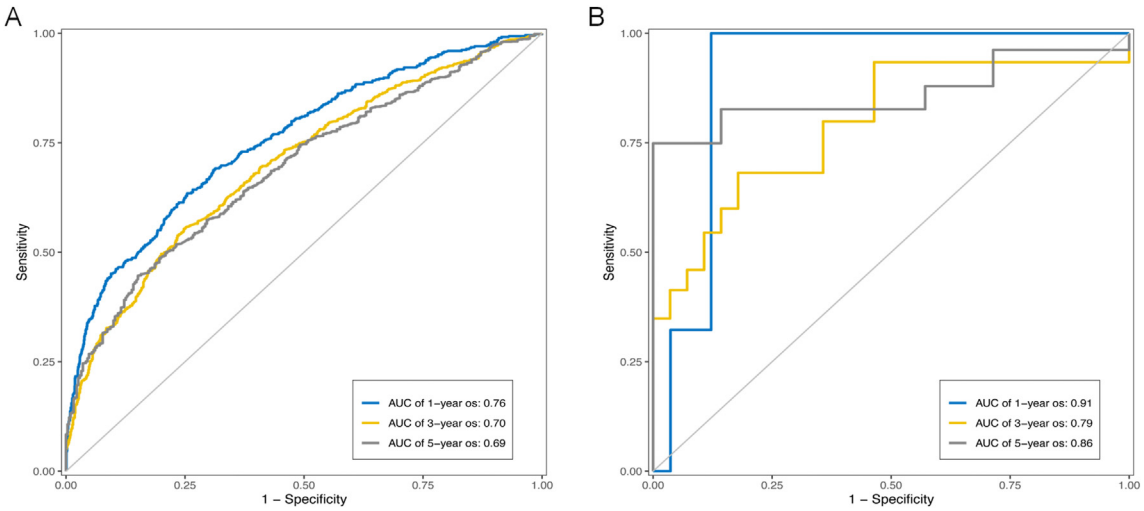


Figure 3. Receiver operating characteristic curves and the area under the time-dependent receiver operating characteristic curve for 1-, 3-, and 5-year overall survival in (A) training and (B) validation cohorts.

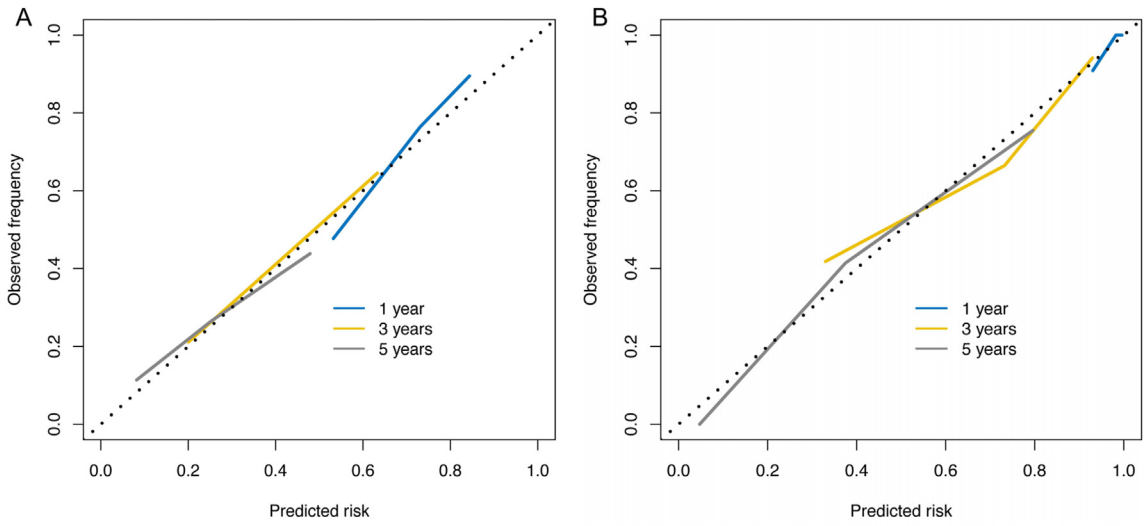


Figure 4. Calibration curves of 1-, 3-, and 5-year overall survival in (A) training and (B) validation cohorts.

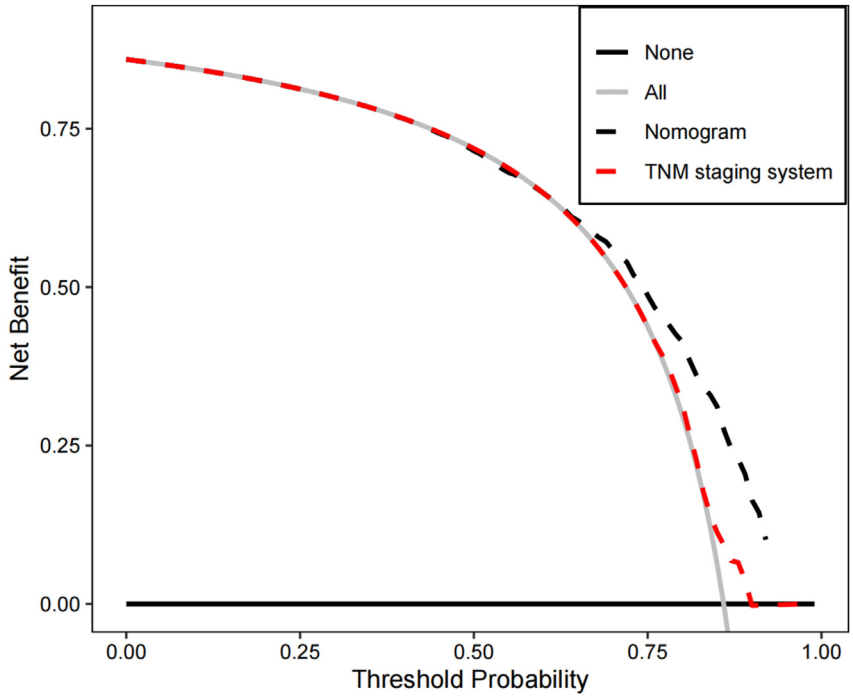


Figure 5. Decision curve analyses of the nomogram and 8th edition AJCC-TNM staging system. AJCC: American Joint Committee on Cancer; TNM: Tumor node metastasis.

survival curves of the patients in each risk group in the training and validation cohorts [Figure 7]. The log-rank test was used to compare the OS of each risk group, and the results showed statistically significant differences ($P < 0.05$), suggesting that this risk stratification can accurately identify groups of patients with *de novo* MBC with different risk factors.

Discussion

This study used data from patients with *de novo* MBC extracted from the SEER database to identify prognostic factors and develop nomograms to predict the 1-, 3-, and 5-year OS. The results were validated in a retrospective study of 103 patients with *de novo* MBC at an institution in northwest China. Both the internal and external validations of the nomogram showed good discriminative ability. In addition, we divided the patients into different risk groups using a nomogram that can predict the prognosis of different risk groups and guide precision medicine.

We discovered that the median survival of patients with *de novo* MBC in the training cohort was 29.0 months, and the 5-year survival rate was 27.3%, which is consistent with that in previous reports.[17] However, the median survival of patients with *de novo* MBC in the validation cohort was 54.0 months, and the 5-year survival rate was 45.3%, which was significantly better than that in the training cohort. This may be related to the latest progress in the diagnosis and treatment of *de novo* MBC, particularly biological therapy. At present, trastuzumab combined with pertuzumab has become the standard treatment for HER2-positive MBC; cyclin-dependent kinase 4/6 (CDK4/6) inhibitors are also widely used in HR+ MBC, which improves the OS of patients with *de novo* MBC.[18] In addition, chemotherapy also has an impact on the prognosis of MBC.[19] In the validation cohort, more than 90% of the patients received

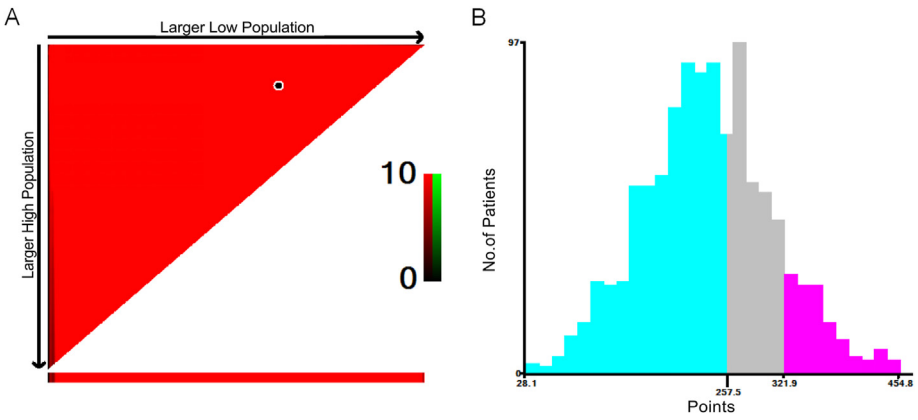


Figure 6. X-tile plots to identify the optimal risk score cutoff based on overall survival (A and B).

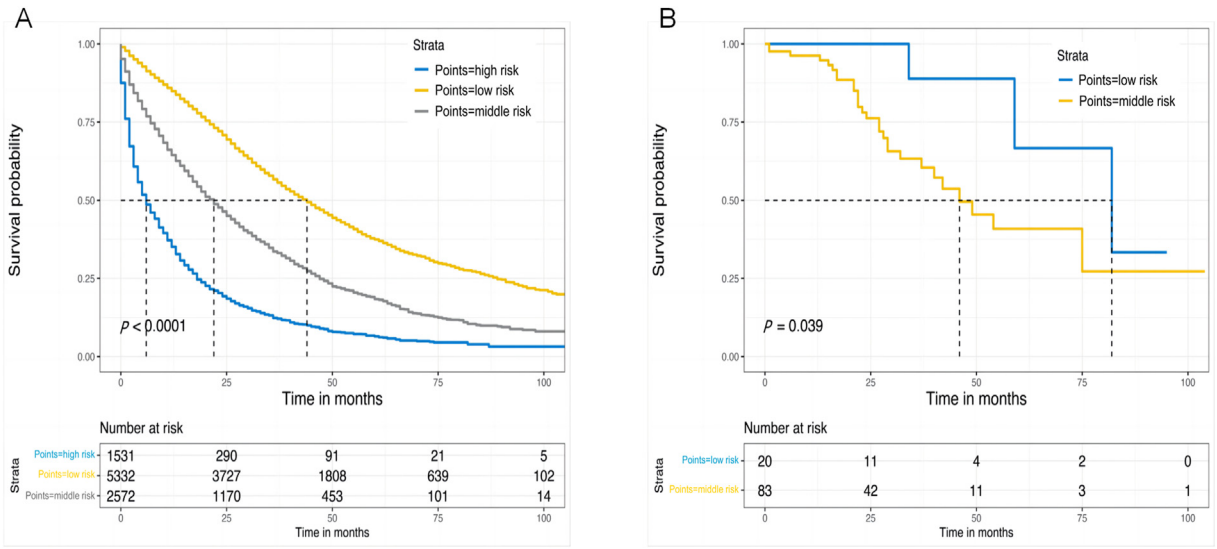


Figure 7. Kaplan–Meier survival curve of overall survival risk subgroup in (A) training cohort and (B) validation cohorts.

chemotherapy, and the proportion of chemotherapy was significantly higher than that in the training cohort (54.1%), which also resulted in a better OS in the validation cohort than in the training cohort.

MBC is a heterogeneous disease and many factors affect its prognosis and drug efficacy. The subtype is a vital prognostic factor for breast cancer and plays an important role in the scoring system.[20,21] As shown in our nomogram, the HR + MBC subtype scored lower and the TNBC subtype scored higher, indicating that the TNBC subtype had the worst prognosis in terms of molecular typing, which was consistent with previous reports.[22] TNBC lacks targeted drugs for HR- and HER2-subtypes, and has a higher risk of recurrence, metastasis, and drug resistance.[23] Additionally, several studies have indicated that patients with TNBC have a relatively high expression of epidermal growth factor receptor (EGFR), which may increase the risk of brain metastasis in patients with cancer. Therefore, the probability of brain metastasis in patients with TNBC is significantly higher than in other types of breast cancer, resulting in a poor prognosis.[24]

It has been reported that the location of distant metastases is associated with the survival of patients with MBC. Patients with bone metastases have the best prognosis, and those with brain and liver metastases have the worst prognosis,[25–27] which is consistent with the scores calculated according to different metastatic sites in our nomogram. Additionally, as shown in our nomogram, the score for surgical treatment was lower, suggesting that surgical treatment is beneficial for the prognosis of patients with MBC, which is consistent with similar findings found in several retrospective studies.[28,29] This may be due to the reduction of tumor burden after local lesion resection, reducing tumor metastasis and recurrence.[30] However, several prospective trials have shown divergent results.[31,32] The results of the 2015 Indian TATA study showed that compared with no local treatment, the local treatment group did not improve the OS (median OS 19.2 months *vs.* 20.5 months).[31] The EA2108 study indicated that surgical treatment can improve local symptoms, but cannot improve the survival rate, and has no effect on the overall quality of life.[32] We anticipate that the JCOG1017 study may address current discrepancies in the literature.[33]

In addition, multivariate Cox analysis showed that age at diagnosis, histological grade, T stage, N stage, chemotherapy, and radiotherapy also affected overall survival. All these predictors were combined to form a nomogram and divide the population into low-, intermediate-, and high-risk groups according to the optimal cutoff values. Different treatment plans were formulated for the groups with different risk levels. No high-risk individuals were identified in the validation cohort. We believe that this is related to improvements in medical practice in recent years, which have reduced the risk of MBC in patients. Another possibility is that the small number of cases in the validation cohort contributed to this result; therefore, a larger sample size, including more related factors, is needed for validation.

The DCA curve results showed that the nomogram developed in our study was superior to the 8th edition AJCC in the risk stratification of patients with *de novo* MBC; therefore, it could better predict the survival

probability of these patients and guide diagnosis and treatment. Patients who are defined as high-risk by the nomogram are recommended to strengthen follow-up and intensive treatment. In addition, multi-gene tests, such as the 21-gene recurrence score, have been gradually used to predict recurrence risk in patients with early-stage HR+ breast cancer.[34] We believe that the combination of nomograms and genomics can better guide the risk of clinical decision-making in patients with *de novo* MBC.

Although this study provides a reference for other studies, it has certain limitations. First, this was a retrospective study with inherent biases in the study design. Second, the SEER database included several other prognostic factors, including metastatic sites other than the bone, lung, liver, and brain; the number and size of metastatic lesions; the Ki67 index; and the use of endocrine and targeted therapies. The nomogram that we created did not include other prognostic factors. Such information would allow for a more specific prognostic assessment of patients with MBC. Third, our study was limited in its ability to assess treatment after recurrence or progression. As the SEER database only provides information on the first course of treatment at the time of diagnosis, we were unable to determine the recurrence risk for patients with *de novo* MBC in our study. Furthermore, our validation cohort included retrospective data from a single medical institution representing only a population in northwest China. Other populations from different countries or prospective studies are required for further external validation of the nomogram. These limitations should be investigated in future studies.

In conclusion, the established nomogram has favorable predictive and discriminative abilities for OS in patients with *de novo* MBC and can be used to stratify patients into different risk subgroups. This may be a useful tool in clinical practice. Other populations from different countries or prospective studies are needed to further validate the nomogram.

Funding

None.

Author contributions

Xinlan Liu and Jiuda Zhao: Conceptualization, methodology, supervision, review, and editing; Lizhi Ning and Yaobang Liu: Methodology, software, data curation, visualization, and writing of original draft preparation; Yujin Hou and Mingqiang Shi: Supervision and review; Miaozhou Wang and Zhen Liu: Supervision and editing.

Ethics statement

The study protocol was reviewed and approved by the Ethics Committee of Ningxia Medical University General Hospital (No. KYLL-2022-1398), and the final approval date was 02 December 2022. Because SEER data are publicly available, this study did not require approval from the institutional ethics committee. The requirement for patient consent was waived because this was a retrospective study that conformed to the provisions of the *Declaration of Helsinki.*

Data availability

The data from SEER are publicly available for use in accordance with a limited use agreement for SEER research data: the Surveillance, Epidemiology, and End Results (SEER) Program (https://seer.cancer.gov), SEER*Stat Database. The data from Ningxia Medical University General Hospital used in the current study are available from the corresponding author on reasonable request.

Conflicts of interest

None.

Acknowledgments

None.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J Clin*. 2021;71:209–249. https://doi.org/10.3322/caac.21660.
2. Liu X, Li J, Cadilha BL, et al. Epithelial-type systemic breast carcinoma cells with a restricted mesenchymal transition are a major source of metastasis. *Sci Adv*. 2019;5: eaav4275. https://doi.org/10.1126/sciadv.aav4275.
3. Sánchez-Muñoz A, Pérez-Ruiz E, Ribelles N, Márquez A, Alba E. Maintenance treatment in metastatic breast cancer. *Expert Rev Anticancer Ther*. 2008;8:1907–1912. https://doi.org/10.1586/14737140.8.12.1907.
4. Si Y, Yuan P, Hu N, et al. Primary tumor surgery for patients with *de novo* stage IV breast cancer can decrease local symptoms and improve quality of life. *Ann Surg Oncol*. 2020;27:1025–1033. https://doi.org/10.1245/s10434-019-08092-2.
5. Lane WO, Thomas SM, Blitzblau RC, et al. Surgical resection of the primary tumor in women with de novo stage IV breast cancer: contemporary practice patterns and survival analysis. *Ann Surg*. 2019;269:537–544. https://doi.org/10.1097/SLA.0000000000002621.
6. Azim HA, Kroman N, Paesmans M, et al. Prognostic impact of pregnancy after breast cancer according to estrogen receptor status: a multicenter retrospective study. *J Clin Oncol*. 2013;31:73–79. https://doi.org/10.1200/JCO.2012.44.2285.
7. Wang Y, Li J, Xia Y, et al. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J Clin Oncol*. 2013;31:1188–1195. https://doi.org/10.1200/JCO.2012.41.5984.
8. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non–small-cell lung cancer. *J Clin Oncol*. 2015;33:861–869. https://doi.org/10.1200/JCO.2014.56.6661.
9. Zhang G, Wu Y, Zhang J, et al. Nomograms for predicting long-term overall survival and disease-specific survival of patients with clear cell renal cell carcinoma. *OncoTargets Ther*. 2018;11:5535–5544. https://doi.org/10.2147/OTT.S171881.
10. El Sharouni MA, Ahmed T, Varey AHR, et al. Development and validation of nomograms to predict local, regional, and distant recurrence in patients with thin (T1) melanomas. *J Clin Oncol*. 2021;39:1243–1252. https://doi.org/10.1200/JCO.20.02446.
11. Lobbezoo DJA, van Kampen RJW, Voogd AC, et al. Prognosis of metastatic breast cancer: are there differences between patients with de novo and recurrent metastatic breast cancer? *Br J Cancer*. 2015;112:1445–1451. https://doi.org/10.1038/bjc.2015.127.
12. Lee CK, Hudson M, Stockler M, et al. A nomogram to predict survival time in women starting first-line chemotherapy for advanced breast cancer. *Breast Cancer Res Treat*. 2011;129:467–476. https://doi.org/10.1007/s10549-011-1471-9.
13. Li S, Zhao J, Zhu L, Su F, Chen K. Development and validation of a nomogram predicting the overall survival of stage IV breast cancer patients. *Cancer Med*. 2017;6: 2586–2594. https://doi.org/10.1002/cam4.1224.
14. Dawood S, Broglio K, Ensor J, Hortobagyi GN, Giordano SH. Survival differences among women with de novo stage IV and relapsed breast cancer. *Ann Oncol*. 2010; 21:2169–2174. https://doi.org/10.1093/annonc/mdq220.
15. Bhoo-Pathy N, Verkooijen HM, Tan EY, et al. Trends in presentation, management, and survival of patients with de novo metastatic breast cancer in a Southeast Asian setting. *Sci Rep*. 2015;5:16252. https://doi.org/10.1038/srep16252.
16. Zhao W, Wu L, Zhao A, et al. A nomogram for predicting survival in patients with de novo metastatic breast cancer: a population-based study. *BMC Cancer*. 2020;20:982. https://doi.org/10.1186/s12885-020-07449-1.
17. Breast Cancer Expert Committee of National Cancer Quality Control Center, Breast cancer expert committee of China anti-cancer association, cancer drug clinical research committee of China anti-cancer association. Guidelines for clinical diagnosis and treatment of advanced breast cancer in China (2022 edition) [In Chinese]. *Chin J Oncol*. 2022;44:1262–1287. https://doi.org/10.3760/cma.j.cn112152-20221007-00680.
18. Breast Cancer Expert Committee of National Cancer Quality Control Center, Cancer Drug Clinical Research Committee of China Anti-Cancer Association. Consensus recommendations for the clinical application of CDK4/6 inhibitors in patients with hormone receptor positive, human epidermal growth factor receptor 2 negative advanced breast cancer [In Chinese]. *Chin J Oncol*. 2021;43:405–413. https://doi.org/10.3760/cma.j.cn112152-20210113-00045.
19. Lin NU, Vanderplas A, Hughes ME, et al. Clinicopathologic features, patterns of recurrence, and survival among women with triple-negative breast cancer in the National Comprehensive Cancer Network. *Cancer*. 2012;118:5463–5472. https://doi.org/10.1002/cncr.27581.
20. Gong Y, Liu YR, Ji P, Hu X, Shao ZM. Impact of molecular subtypes on metastatic breast cancer patients: a SEER population-based study. *Sci Rep*. 2017;7:45411. https://doi.org/10.1038/srep45411.
21. Lee J, Park S, Kim S, et al. Characteristics and survival of breast cancer patients with multiple synchronous or metachronous primary cancers. *Yonsei Med J*. 2015;56: 1213. https://doi.org/10.3349/ymj.2015.56.5.1213.
22. Schettini F, Martínez-Sáez O, Falato C, et al. Prognostic value of intrinsic subtypes in hormone-receptor-positive metastatic breast cancer: systematic review and meta-

analysis. *ESMO Open*. 2023;8:101214. https://doi.org/10.1016/j.esmoop.2023.101214.
23. Xie Q, Yang Z, Huang X, et al. Ilamycin C induces apoptosis and inhibits migration and invasion in triple-negative breast cancer by suppressing IL-6/STAT3 pathway. *J Hematol Oncol*. 2019;12:60. https://doi.org/10.1186/s13045-019-0744-3.
24. Wang H, Zhang C, Zhang J, Kong L, Zhu H, Yu J. The prognosis analysis of different metastasis pattern in patients with different breast cancer subtypes: a SEER based study. *Oncotarget*. 2017;8:26368–26379. https://doi.org/10.18632/oncotarget.14300.
25. Ording AG, Heide-Jørgensen U, Christiansen CF, Nørgaard M, Acquavella J, Sørensen HT. Site of metastasis and breast cancer mortality: a Danish nationwide registry-based cohort study. *Clin Exp Metastasis*. 2017;34:93–101. https://doi.org/10.1007/s10585-016-9824-8.
26. Duan XF, Dong NN, Zhang T, Li Q. The prognostic analysis of clinical breast cancer subtypes among patients with liver metastases from breast cancer. *Int J Clin Oncol*. 2013;18:26–32. https://doi.org/10.1007/s10147-011-0336-x.
27. Lee ES, Jung SY, Kim JY, et al. Identifying the potential long-term survivors among breast cancer patients with distant metastasis. *Ann Oncol*. 2016;27:828–833. https://doi.org/10.1093/annonc/mdw036.
28. Desille-Gbaguidi H, Avigdor S, Body G, Ouldamer L. Survival impact of primary site surgery on metastatic breast cancer patients at diagnosis. *J Gynecol Obstet Hum Reprod*. 2019;48:171–177. https://doi.org/10.1016/j.jogoh.2018.10.014.
29. Vohra NA, Brinkley J, Kachare S, Muzaffar M. Primary tumor resection in metastatic breast cancer: a propensity-matched analysis, 1988-2011 SEER data base. *Breast J*. 2018;24:549–554. https://doi.org/10.1111/tbj.13005.
30. 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;28:5059–5070. https://doi.org/10.1245/s10434-021-09650-3.
31. Badwe R, Hawaldar R, Nair N, et al. Locoregional treatment versus no treatment of the primary tumour in metastatic breast cancer: an open-label randomised controlled trial. *Lancet Oncol*. 2015;16:1380–1388. https://doi.org/10.1016/S1470-2045(15)00135-7.
32. Ulaş Kahya B, Sütcüoğlu O, Yazıcı O, Özdemir N. Early local therapy for the primary site in de novo stage IV breast cancer: results of a randomized clinical trial (EA2108). *J Clin Oncol*. 2022;40:3898. https://doi.org/10.1200/JCO.22.00404.
33. Shien T, Mizutani T, Tanaka K, et al. A randomized controlled trial comparing primary tumor resection plus systemic therapy with systemic therapy alone in metastatic breast cancer (PRIM-BC): Japan Clinical Oncology Group Study JCOG1017. *J Clin Oncol*. 2017;35:TPS588. https://doi.org/10.1200/JCO.2017.35.15_suppl.TPS588.
34. Wang C, Xu Y, Lin Y, et al. Comparison of CTS5 risk model and 21-gene recurrence score assay in large-scale breast cancer population and combination of CTS5 and recurrence score to develop a novel nomogram for prognosis prediction. *Breast*. 2022;63:61–70. https://doi.org/10.1016/j.breast.2022.03.007.