# Nomogram for predicting distant metastasis of male breast cancer
## A SEER population-based study

Dasong Wang, MD[a], Lei Yang, MD[a], Yan Yang, MD[a], Maoshan Chen, MD[a], Hongwei Yang, MD[a],*

## Abstract

The main purpose of this study was to build a prediction model for male breast cancer (MBC) patients to predict the possibility of distant metastasis. The Surveillance, Epidemiology, and End Results database was used to obtain data on patients with MBC. The patients were divided into a training set and a validation set at a ratio of 7:3. The risk variables of distant metastasis in the training set were determined by univariate and multivariate logistic regression analyses. And then we integrated those risk factors to construct the nomogram. The prediction nomogram was further verified in the verification set. The discrimination and calibration of the nomogram were evaluated by the area under the receiver operating characteristic curve, calibration plots, respectively. A total of 1974 patients (1381 in training set and 593 in validation set) were eligible for final inclusion, of whom 149 (7.55%) had distant metastasis at the diagnosed time. Multivariate logistic regression analyses presented that age, T stage, N stage, and hormone receptor status were independent risk factors for distant metastasis at initial diagnosis of male breast cancer. Finally, the 4 variables were combined to construct the nomogram. The area under the curve values for the nomogram established in the training set and validation set were 0.8224 (95%CI: 0.7796–0.8652) and 0.8631 (95%CI: 0.7937–0.9326), suggesting that the nomogram had good predictive power. The calibration plots illustrated an acceptable correlation between the prediction by nomogram and the actual observation, as the calibration curve was closed to the diagonal bisector line. An easy-to-use nomogram, being proven to be with reliable discrimination ability and accuracy, was established to predict distant metastasis for male patients with breast cancer using the easily available risk factors.

**Abbreviations:** ER = estrogen receptor, HER2 = human epidermal growth factor receptor-2, HR = hormone receptor, MBC = male breast cancer, PR = progesterone receptor, SEER = Surveillance, Epidemiology, and End Results Program.

**Keywords:** distant metastasis, male breast cancer, nomogram, SEER

## 1. Introduction

Male breast cancer (MBC) is a relatively rare malignant tumor, accounting for less than 1% of all male malignant tumors and about 1% of all breast cancers, but the incidence has gradually increased in recent years.[1–3] Due to the low incidence of MBC, it has been excluded from most prospective clinical randomized controlled studies, resulting in a lack of prospective data and guidelines for MBC. Most treatment regimens about MBC are based on data generated by female patients or on data with lower level of evidences (such as small retrospective cohort studies or case reports).[2,4] However, there are significant differences in clinicopathologic characteristics and prognosis between male and female breast cancer patients. MBC patients usually present with larger, more advanced tumors than women, along with lymph node metastasis.[5,6] MBC patients are usually older when diagnosed and receive less standardized treatment.[7] Men with breast cancer have a worse prognosis than women, with 5-year and overall survival

rates about 10% and 15% lower than women, respectively. Differences in clinicopathologic characteristics and treatment were the most common factors contributing to differences in prognosis, but the differences in prognosis still persisted after adjusting for age, race, and treatment.[7] Guiding the treatment of MBC patients by referring to the researches about female patients has certain limitations, so it is necessary to conduct research on MBC.

Distant metastasis is an important cause of death in patients with breast cancer at present. Although the survival rate of patients with metastatic breast cancer has improved in the past decades, it still cannot be completely cured.[8–10] As far as we can concerned, there are few studies on the relationship between distant metastasis of breast cancer and biological factors, and there are no consistent conclusions on the risk factors of distant metastasis of breast cancer.[11–13] Understanding the risk factors of distant metastasis in MBC is helpful in guiding clinical diagnosis and treatment, and improving the prognosis of MBC. The purpose of this study is to retrospectively analyze

the clinicopathologic characteristics of MBC by using the data of SEER database, and to explore the possible relationship between primary tumor and distant metastasis, so as to guide the individual diagnosis and treatment of patients.

## 2. Data and Methods

### 2.1. Data source

We obtained data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program. SEER currently collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 30% of the U.S. population.[14,15] SEER began collecting sites of metastasis at initial diagnosis in 2010; because of this, we used 2010 as the starting point for our study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All information from the SEER program is available and free for public, so the agreement of the medical ethics committee board was not necessary.

### 2.2. Patient selection

We extracted all cases of men with microscopically confirmed breast cancer at initial presentation, diagnosed between January 2010 and December 2016. Inclusion criteria: Pathology diagnosed with invasive breast cancer; male; the diagnosis was made between January 2010 and December 2016. Exclusion criteria: Clinical or pathological information is incomplete; not the only primary malignancy; non-unilateral breast cancer; breast cancer

was diagnosed at autopsy. Patient screening flow chart is shown in Figure 1.

### 2.3. Statistical analysis

According to inclusion criteria and exclusion criteria, a total of 1974 MBC patients were screened for analysis. Study variables included age, race, tumor subtype, pathological grade, sites of metastases, tumor stage, estrogen receptor (ER) status, progesterone receptor (PR) status, human epidermal growth factor receptor-2 (HER2) status, and other factors. The main research index was the occurrence of distant metastasis. In SEER database, distant metastasis status was defined as the occurrence of distant metastasis at the time of first diagnosis of malignant tumor. In this study, metastasis of bone, liver, lung, brain was mainly included. Since there is no causal relationship between surgical treatment, radiotherapy, chemotherapy, and other research indicators and the main research indicators of this study, the above indicators are not included in the further analysis. Four tumor grades were collapsed into 3 grades, with grade 3 and 4 merged. Tumor stage was registered according to the American Joint Committee on Cancer staging system, 7th edition.

All included patients were randomly divided into the training set (1381 cases) and the validation set (593 cases) in a 7:3 ratio. All data were statistically analyzed and plotted using IBM SPSS 23.0 or Stata15.0 software. Chi-square ($\chi$2) test was used for comparison of counting data, and $t$ test was used for comparison of measurement data. Univariate logistic regression was used to identify risk factors associated with the incidence of distant metastasis in the training group, and statistically significant
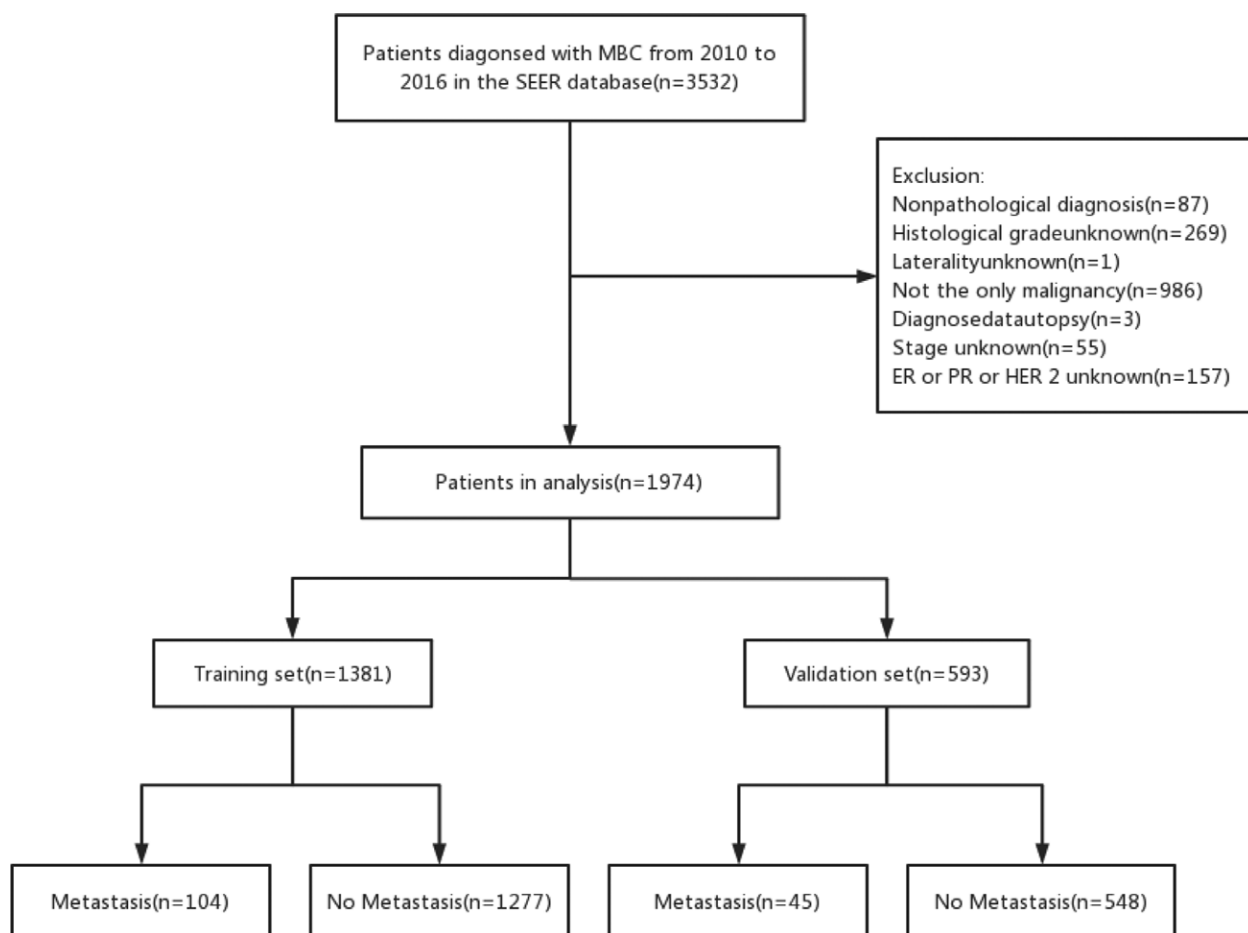


**Figure 1.** The flowchart of this study.

variables were included in multivariate logistic regression analysis. $P < .05$ was considered statistically significant. Stata15.0 software was used to construct the prognostic model, and the results were visualized. Receiver operating characteristic curves and calibration curves were used to verify the accuracy of the model.

## 3. Results

### 3.1. Clinicopathologic characteristics of MBC

In this study, a total of 1974 patients with MBC diagnosed between 2010 and 2016 were included, of whom 149 (7.55%) had distant metastasis and 1825 (92.45%) had no distant metastasis, with rates similar to those reported in previous studies.[16,17] The results showed that the average age of the included patients was 65.03 years old (range 22–99 years old), and the main onset age was 60–79 years old (elderly group), accounting for 53.55% of all patients. The majority of patients were white (78.52%), and the most common pathologic type was invasive ductal carcinoma (84.59%). Histological grade I and II accounted for 64.23%, stage T1 and T2 staging accounted for 42.6% and 43.67%, respectively. No axillary lymph node metastasis accounted for 53.5%. 96.96% of patients were ER positive, 90.27% were PR positive, and 13.27% were HER2 positive. The most common tumor subtype was hormone receptor (HR)+/HER2–, which accounted for 84.60% of all patients. There were no significant differences in clinicopathologic characteristics between the training set and the validation set, and the specific clinicopathologic characteristics are shown in Table 1.

### 3.2. Factors associated with distant metastasis

The logistic regression model was established to evaluate the clinicopathologic characteristics associated with distant metastasis. Univariate logistic regression analysis of the training set showed that age, race, marital status, T stage, N stage, HR status, HER2 status were significantly related to distant metastasis ($P < .05$). ER status and PR status have similar guiding value for endocrine therapy in MBC patients. Therefore, ER or PR positive patients were defined as HR positive in the study. Unmarried, divorced, widowed, and separated marital status were all considered to be adverse factors affecting prognosis,[18] and the sample size was small, so they were included in the same group for analysis. There was no correlation between surgical treatment, radiotherapy, chemotherapy, and distant metastasis of MBC patients at initial diagnosis, so they were not included in logistic regression analysis.

Variables with statistical significance in univariate logistic regression analysis such as age and race were included in multivariate logistic regression analysis, and the results showed that age, T stage, N stage, and HR status were independent risk factors for distant metastasis at initial diagnosis of MBC. The results of univariate and multivariate logistic regression analysis of the training set are shown in Table 2.

### 3.3. Nomogram construction and validation

A nomogram to predict distant metastasis was established in the training set. Binary logistic regression analyses indicated that age, T stage, N stage, HR status were independent predictive factors of distant metastasis in MBC patients. Therefore, we integrated those variables to construct the nomogram (Fig. 2). According to the nomogram, we can get the score corresponding to each predictor, and the total score corresponds to the predicted probability of distant metastasis.

We drew the receiver operating characteristic curves of predicted probability and calculated the area under the curve values

in the training and validation set (Figs. 3 and 4). The area under the curve values for the nomogram established in the training set and validation set were 0.8224 (95%CI: 0.7796–0.8652) and 0.8631 (95%CI: 0.7937–0.9326), suggesting that the nomogram had good predictive power. We used Stata15.0 software to draw calibration plots of training set and verification set (Figs. 5 and 6). The 95% CI of calibration belt in both training and validation sets did not cross the diagonal bisector line. Therefore, the calibration plot showed good agreement between prediction and observation in the probability of distant metastases.

## 4. Discussion

The survival benefit of cancer patients comes from early detection and standardized treatment, while the occurrence of distant metastasis is an important factor affecting the prognosis of patients. Early prediction or detection of distant metastasis and timely treatment can improve the prognosis of patients.[14,19,20] Breast cancer patients with or without distant metastasis have significantly different prognosis. The 5-year survival rates of local or regional breast cancer patients who were detected early and having actively standardized and systematic treatment were 98.7% and 85.5%, respectively, while the 5-year survival rate of breast cancer patients with distant metastasis was only about 27%.[14] Therefore, distant metastasis is a significant negative prognostic factor for breast cancer, which is crucial for the treatment decision of breast cancer patients. At present, the diagnosis of distant metastasis of breast cancer mainly depends on imaging detection and pathological biopsy, the screening procedure is complicated with high time and economic cost, and it is difficult to diagnose some small metastatic lesions. While nomogram can be used to estimate the distantmetastasis risk by integrating the existing clinicopathological factors, improving diagnostic efficiency, reducing time, and economic cost.[21,22] So it is necessary to establish a simple and sensitive prediction model for distant metastasis in MBC patients. The study established a risk prediction model based on the existing evidence to assess the risk of distant metastasis and screen high-risk groups, so as to guide diagnosis and treatment.

In our study, we found that age, T stage, N stage, HR status were independent predictive factors of distant metastasis in MBC patient, so we integrated those variables to construct the nomogram. Through calibration, it is found that the prediction model has good forecasting ability and accuracy, so it can be used to predict the risk of distant metastasis in MBC patients with relative accuracy.

As far as we can know, there are few studies on the risk factors for distant metastasis of MBC, and more studies have focused on risk factors in female breast cancer. Our study found that risk factors for distant metastasis in MBC included age, T stage, N stage, and HR status, which is consistent with the results of reported studies on female breast cancer patients. Purushotham et al found that age was an independent risk factor for distant metastasis in breast cancer patients. Compared with younger patients, the risk of distant metastasis to bone or viscera was significantly reduced in women of all age groups over 40 years old.[23] Li-jun and James found that age was also a significant predictor of bone metastasis in different predictive models, and younger patients were more likely to develop bone metastasis.[12,24] This is consistent with the results of this study. In our study, the incidence of distant metastasis was 5.65% in the super elderly group (>80 years old), 6.43% in the elderly group (60–79 years old), 9.40% in the middle-aged group (40–59 years old), and 16.33% in the young patients (<40 years old). The probability of distant metastasis increased significantly with decreasing age in the group ($P = .009$).

HR status is one of the most important independent predictors of distant metastasis in breast cancer and has been reported in previous studies. Purushotham et al retrospectively

## Table 1

**Characteristics between the training set and the validation set.**

| Characteristics | All patientsn = 1974(%) | Trainingn = 1381(%) | Validationn = 593(%) | *P* value |
|---|---|---|---|---|
| Age(years) | | | | .973 |
| <40 | 49(2.48) | 34(2.46) | 15(2.53) | |
| 40–59 | 638(32.32) | 445(32.22) | 193(32.55) | |
| 60–79 | 1057(53.55) | 738(53.44) | 319(53.79) | |
| ≥80 | 230(11.65) | 164(11.88) | 66(11.13) | |
| Race | | | | .891 |
| White | 1550(78.52) | 1086(78.64) | 464(78.25) | |
| Black | 289(14.64) | 203(14.70) | 86(14.50) | |
| Others* | 135(6.84) | 92(6.66) | 43(7.25) | |
| Marital status | | | | .170 |
| Married | 1276(64.64) | 885(64.08) | 391(65.94) | |
| Unmarried | 320(16.21) | 240(17.38) | 80(13.49) | |
| Divorced/widowed/separated | 267(13.53) | 180(13.03) | 87(14.67) | |
| Unknown | 111(5.62) | 76(5.5) | 35(5.90) | |
| Laterality | | | | .803 |
| Left | 1057(53.55) | 742(54.73) | 315(53.12) | |
| Right | 917(46.45) | 639(46.27) | 278(46.88) | |
| Tumor location | | | | .273 |
| Central region† | 915(46.35) | 629(45.55) | 286(48.23) | |
| Other region‡ | 1059(53.65) | 752(54.45) | 307(51.77) | |
| Tumor size | | | | .052 |
| <20mm | 855(43.31) | 614(44.46) | 241(40.64) | |
| 20–50mm | 954(48.33) | 664(48.08) | 290(48.90) | |
| >50mm | 165(8.36) | 103(7.46) | 62(10.46) | |
| Grade | | | | .600 |
| I | 223(11.30) | 153(11.08) | 70(11.80) | |
| II | 1033(52.33) | 716(51.85) | 317(53.46) | |
| III/IV | 718(36.37) | 512(37.07) | 206(34.74) | |
| Stage | | | | .155 |
| I | 626(31.71) | 459(33.24) | 167(28.16) | |
| II | 852(43.16) | 584(42.29) | 268(45.19) | |
| III | 347(17.58) | 234(16.94) | 113(19.06) | |
| IV | 149(7.55) | 104(7.53) | 45(7.59) | |
| T | | | | .110 |
| T1 | 841(42.60) | 606(43.88) | 234(39.63) | |
| T2 | 862(43.67) | 597(43.23) | 265(44.69) | |
| T3 | 79(4.00) | 48(3.48) | 31(5.23) | |
| T4 | 192(9.73) | 130(9.41) | 62(10.46) | |
| N | | | | .217 |
| N0 | 1056(53.50) | 762(55.18) | 294(49.58) | |
| N1 | 624(31.61) | 425(30.77) | 199(33.56) | |
| N2 | 183(9.27) | 120(8.69) | 63(10.62) | |
| N3 | 111(5.62) | 69(5.36) | 37(6.24) | |
| M | | | | .964 |
| M0 | 1825(92.45) | 1277(92.47) | 548(92.41) | |
| M1 | 149(7.55) | 104(7.53) | 45(7.59) | |
| Tumor subtype | | | | .872 |
| HR+/HER2− | 1670(84.60) | 1174(85.01) | 496(83.64) | |
| HR+/HER2+ | 248(12.56) | 168(12.17) | 80(13.49) | |
| HR−/HER2+ | 14(0.71) | 10(0.72) | 4(0.67) | |
| HR−/HER2− | 42(2.13) | 29(2.10) | 13(2.19) | |
| ER | | | | .572 |
| Positive | 1914(96.96) | 1314(97.10) | 573(96.63) | |
| Negative | 60(3.04) | 40(2.90) | 20(3.37) | |
| PR | | | | .700 |
| Positive | 1782(90.27) | 1249(90.44) | 533(89.88) | |
| Negative | 192(9.73) | 132(9.56) | 60(10.12) | |
| HER2 | | | | .444 |
| Positive | 262(13.27) | 178(12.89) | 84(14.17) | |
| Negative | 1712(86.73) | 1203(87.11) | 509(85.83) | |
| Distant metastasis | | | | .964 |
| Yes | 149(7.55) | 104(7.53) | 45(7.59) | |
| No | 1825(92.45) | 1277(92.47) | 548(92.41) | |
| Bone metastasis | | | | .653 |
| Yes | 116(5.88) | 79(5.72) | 37(6.24) | |
| No | 1858(94.12) | 1302(94.28) | 556(93.76) | |
| Liver metastasis | | | | .100 |
| Yes | 12(0.61) | 11(0.80) | 1(0.17) | |
| No | 1962(99.39) | 1370(99.20) | 592(99.83) | |

*(Continued)*

**Table1**

*(Continued)*

| Characteristics | All patientsn = 1974(%) | Trainingn = 1381(%) | Validationn = 593(%) | *P* value |
|---|---|---|---|---|
| Lung metastasis | | | | .116 |
| Yes | 54(2.74) | 43(3.11) | 11(1.85) | |
| No | 1920(97.26) | 1338(96.89) | 582(98.15) | |
| Brain metastasis | | | | .491 |
| Yes | 10(0.51) | 6(0.43) | 4(0.67) | |
| No | 1964(99.49) | 1375(99.57) | 589(99.33) | |
| Surgery | | | | .665 |
| No | 178(9.02) | 122(8.83) | 56(9.44) | |
| Yes | 1796(90.98) | 1259(91.17) | 537(90.56) | |
| Radiotherapy | | | | .988 |
| Yes | 583(29.53) | 408(29.54) | 175(29.51) | |
| No | 1391(70.47) | 973(70.76) | 418(70.49) | |
| Chemotherapy | | | | .161 |
| Yes | 822(41.64) | 561(40.62) | 261(44.01) | |
| No | 1152(58.36) | 820(59.38) | 332(55.99) | |

*Includes: American Indian, native Alaskan and Asian, Pacific Islander.

†Tumor was located behind the nipple or areola.

‡Tumor was located in the inner, outer, upper, lower quadrant, axillary tail, or overlapping areas.

**Table 2**

**Univariate and multivariate logistic regression analysis of the training set.**

| Characteristics | Univariate analysis | | Multivariate analysis | |
|---|---|---|---|---|
| | OR(95%CI) | *P* | OR(95%CI) | *P* |
| Age (years) | | | | |
| ≤40 | 4.179(1.347–12.97) | .013 | 3.301(1.089–10.00) | .035 |
| 41–60 | 1.979(0.907–4.316) | .086 | 1.828(0.871–3.835) | .111 |
| 61–80 | 1.387(0.644–2.987) | .404 | 1.406(0.682–2.902) | .356 |
| >80 | Reference | | | |
| Race | | | | |
| White | Reference | | Reference | |
| Black | 1.889(1.159–3.078) | .011 | 1.213(0.687–2.140) | .505 |
| Others* | 1.341(0.625–2.878) | .451 | 1.140(0.471–2.760) | .772 |
| Marital status | | | | |
| Married | Reference | | Reference | |
| Not married† | 1.586(1.061–2.371) | .025 | 1.267(0.800–2.007) | .313 |
| Laterality | | | | |
| Left | Reference | | | |
| Right | 1.082(0.725–1.614) | .701 | | |
| Tumor location | | | | |
| Central region‡ | Reference | | | |
| Other regions§ | 1.401(0.981–2.002) | .064 | | |
| Grade | | | | |
| I | Reference | | | |
| II | 1.499(0.665–3.379) | .329 | | |
| III/IV | 2.207(0.979–4.979) | .056 | | |
| T | | | | |
| T1 | Reference | | Reference | |
| T2 | 4.166(2.060–8.424) | .000 | 2.926(1.551–5.522) | .001 |
| T3 | 19.87(8.044–49.07) | .000 | 17.43(8.113–37.45) | .000 |
| T4 | 25.24(12.09–52.72) | .000 | 18.55(9.471–36.32) | .000 |
| N | | | | |
| N0 | Reference | | Reference | |
| N1 | 4.091(2.451–6.828) | .000 | 2.387(1.467–3.885) | .000 |
| N2 | 4.590(2.321–9.077) | .000 | 2.776(1.504–5.123) | .001 |
| N3 | 7.459(3.586–15.51) | .000 | 3.227(1.597–6.519) | .001 |
| HR | | | | |
| Positive | Reference | | Reference | |
| Negative | 4.870(2.628–9.027) | .000 | 6.782(3.102–14.83) | .000 |
| HER2 | | | | |
| Positive | Reference | | Reference | |
| Negative | 0.430(0.266–0.695) | .001 | 0.660(0.381–1.145) | .139 |

*Includes: American Indian, native Alaskan and Asian, Pacific Islander.

†Includes: Unmarried, divorced, widowed, and separated.

‡Tumor was located behind the nipple or areola.

§Tumor was located in the inner, outer, upper, lower quadrant, axillary tail, or overlapping areas.
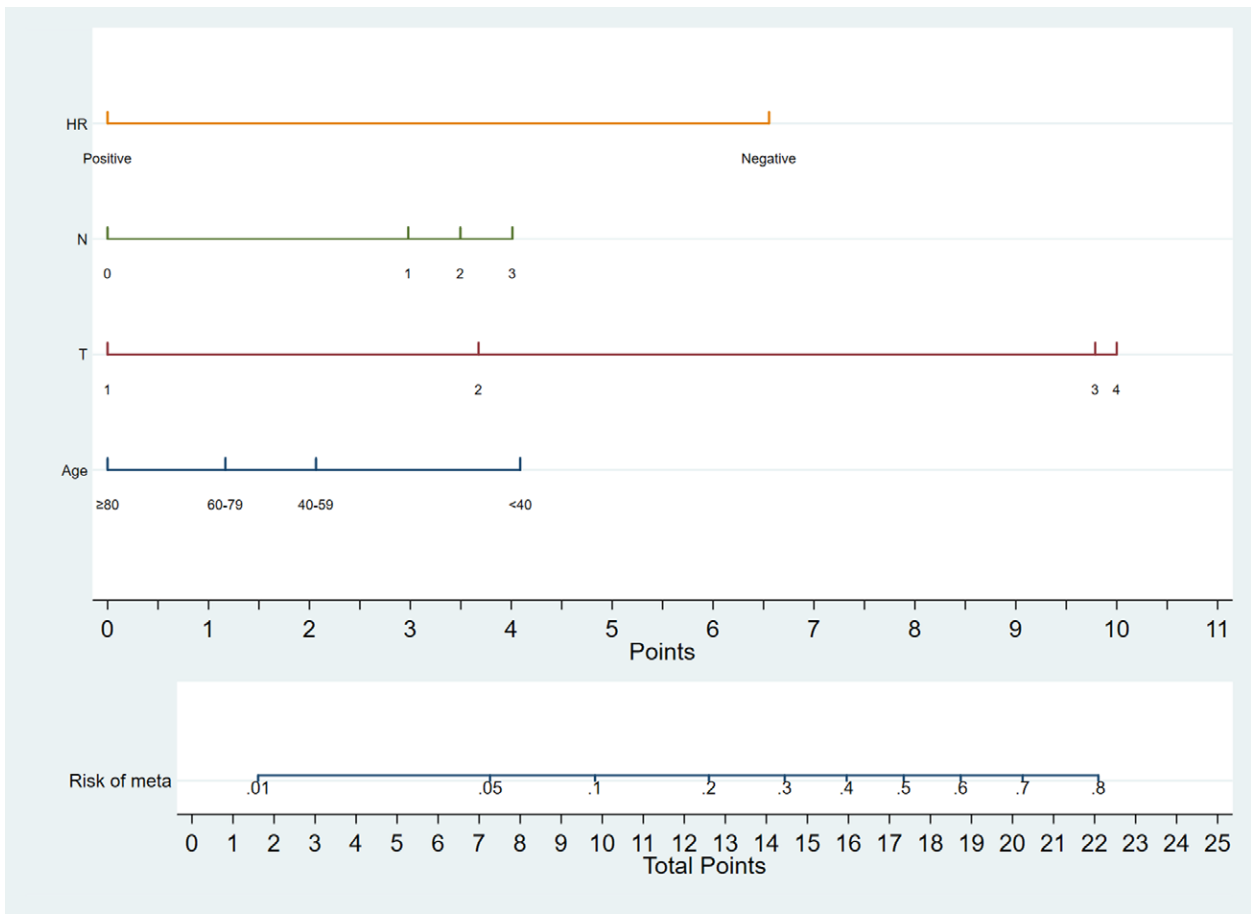
**Figure 2.** Nomograms for predicting the risk of distant metastasis in male patients with breast cancer.
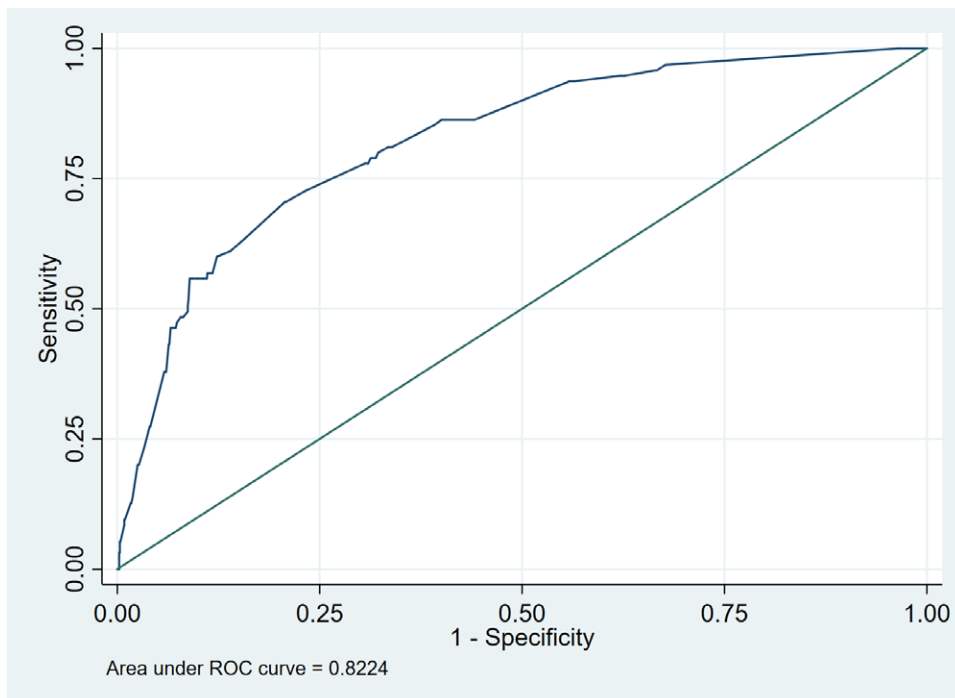


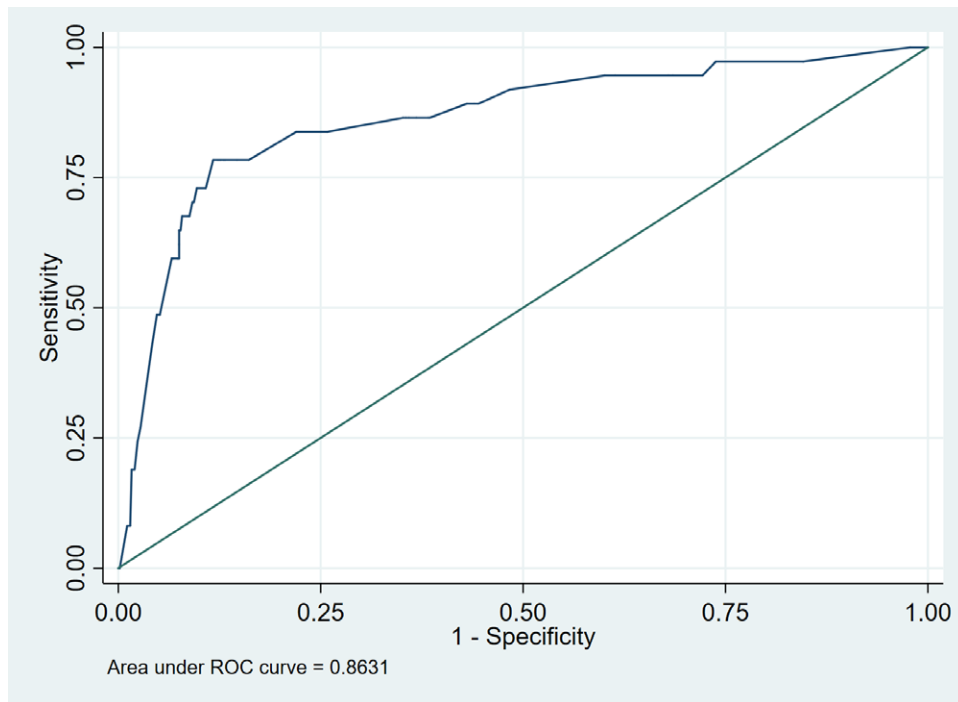**Figure 3.** Receiver operating characteristic (ROC) analyses of the nomogram in training set.

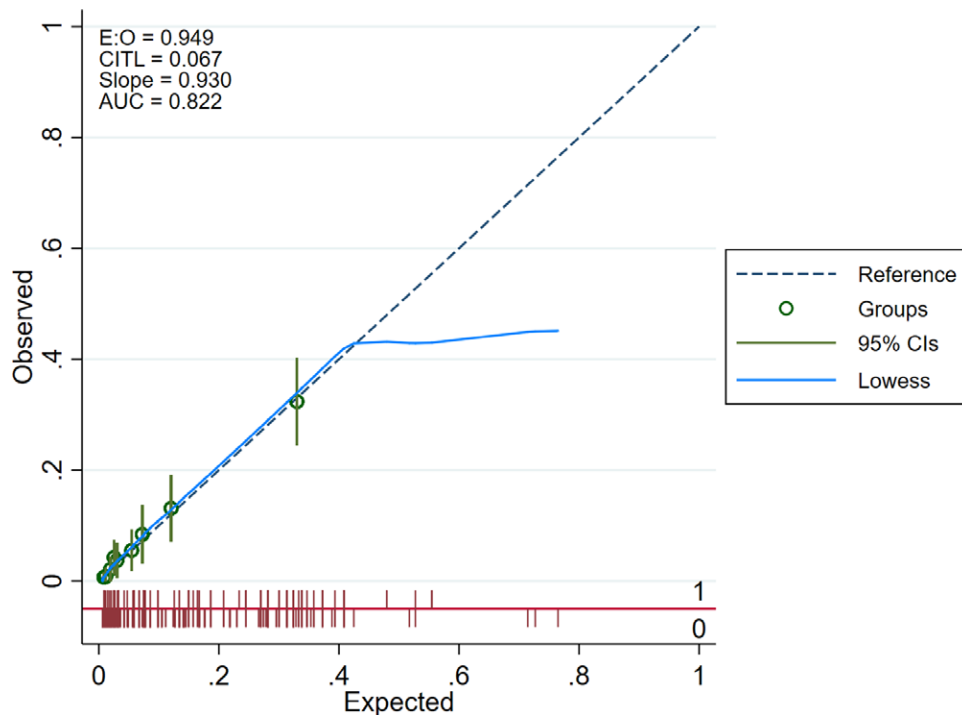**Figure 4.** Receiver operating characteristic (ROC) analyses of the nomogram in validation set.



**Figure 5.** Calibration plots of the nomogram in training set.

analyzed the medical records of 3553 BC patients in Guy's Hospital from January 1, 1986 to December 30, 2006, and showed that patients with ER positive are less likely to develop distant metastasis than patients with ER negative (HR 0.74, 95%CI: 0.61–0.90). And they also found that patients with PR positive are less likely to develop distant metastasis than patients with PR negative (HR: 0.77, 95%CI: 0.51–0.83).[23] Zhenhai et al retrospectively analyzed the data of 6238

patients with breast cancer in the SEER database from 2010 to 2013 and found that ER negative and PR negative were both significant predictors of liver metastasis in breast cancer patients.[25] In this study, we found that the probability of distant metastasis was 6.99% in HR positive MBC, 26.79% in HR negative MBC ($P < .001$), the risk of distant metastasis was significantly higher in patients with HR negative than in patients with HR positive.
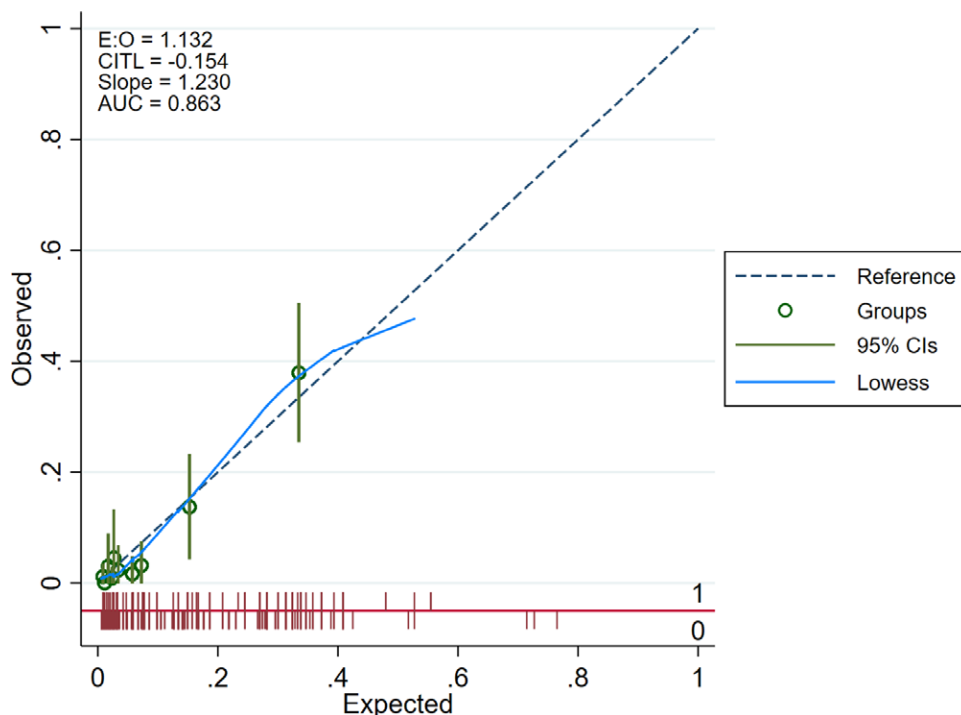
**Figure 6.** Calibration plots of the nomogram in validation set.

The local tumor status is also a predictor of distant metastasis in breast cancer patients. Rudolph et al found that tumor diameter, histological grade, lymph node status, Ki-67, and PR status were independent predictors of distant metastasis in breast cancer.[26] Bozcuk et al found that tumor diameter was associated with distant metastasis in breast cancer patients, and the risk of distant metastasis increased with the increase of tumor diameter.[27] This finding was confirmed by Purushotham et al They found that the risk of distant metastasis was 1.61 (95%CI 1.37–1.89) times higher in patients with tumor size of 2 to 5 cm than in patients with tumor size of <2 cm, and 2.35 (95%CI 1.76–3.14) times higher in patients with tumor size of >5 cm than in patients with tumor size of <2 cm.[23] Purushotham et al also found that lymph node metastasis was an independent predictor of distant metastasis in breast cancer patients. Patients with 1 to 3 axillary lymph node metastasis had a higher risk of distant metastasis than patients without axillary lymph node metastasis (HR 2.09, 95%CI: 1.71–2.56). And the risk was further increased when patients had 4 or more than 4 axillary lymph node metastasis (HR 5.11, 95%CI: 4.14–6.30).[23] In this study, the probability of distant metastasis was 1.55% in T1 patients, 5.68% in T2 patients, 26.58% in T3 patients, and 34.38% in T4 patients. With the increase of T stage, the risk of distant metastasis in MBC patients increased significantly ($P < .001$). The incidence of distant metastasis was 2.94% in patients with N0 stage, 10.58% in N1 patients, 14.21% in N2 patients, and 23.42 in N3 patients. With the increase of N stage, the risk of distant metastasis in MBC patients increased significantly ($P < .001$).

In this study, we established a clinical prediction model for distant metastasis of MBC based on a large sample of population data from the SEER database. The established model is based on readily available clinicopathological factors and has good predictive power and accuracy. Therefore, this model can be used in clinical application to assess the risk of distant metastasis in MBC patients, so as to specify an individual examination and treatment plan.

## 5. Limitations

Although our model has good accuracy in predicting distant metastasis of MBC, there are still some potential limitations to our study. This study is a retrospective study, so there may be selection bias in the included cases. For example, some cases may be excluded from this study due to partial information deficiency. Secondly, some predictors that may be associated with distant metastasis (such as KI-67, etc.) could not be analyzed due to the absence of this indicator in the SEER database. Finally, on account of the data for both the training set and the validation set come from the SEER database, this may lead to overfitting of the model, so it needs to be validated in other external data to prove its reproducibility.

## 6. Conclusion

In conclusion, through logistic regression analysis, we found that age, T stage, N stage, HR status were independent risk factors associated with distant metastasis in MBC. Based on these clinical risk factors, we established the nomogram that could accurately and easily predict the distant metastasis risk of MBC patients, thus contributing to treatment decision making.

## Author contributions

All authors contributed to the study conception and design. Data collection and analysis were performed by DW. The first draft of the manuscript was written by DW, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.
**Conceptualization:** Dasong Wang, Hongwei Yang.
**Data curation:** Dasong Wang, Lei Yang.
**Formal analysis:** Dasong Wang, Lei Yang, Yan Yang, Maoshan Chen, Hongwei Yang.
**Investigation:** Dasong Wang, Yan Yang, Maoshan Chen.
**Methodology:** Dasong Wang, Lei Yang, Yan Yang, Maoshan Chen.

## References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69:7–34.

[2] Leon-Ferre RA, Giridhar KV, Hieken TJ, et al. A contemporary review of male breast cancer: current evidence and unanswered questions. Cancer Metastasis Rev. 2018;37:599–614.

[3] Yalaza M, Inan A, Bozer M. Male breast cancer. J Breast Health. 2016;12:1–8.

[4] Duma N, Hoversten KP, Ruddy KJ. Exclusion of male patients in breast cancer clinical trials. JNCI Cancer Spectr. 2018;2:pky018.

[5] Miao H, Verkooijen HM, Chia KS, et al. Incidence and outcome of male breast cancer: an international population-based study. J Clin Oncol. 2011;29:4381–6.

[6] Anderson WF, Jatoi I, Tse J, et al. Male breast cancer: a population-based comparison with female breast cancer. J Clin Oncol. 2010;28:232–9.

[7] Wang F, Shu X, Meszoely I, et al. Overall mortality after diagnosis of breast cancer in men vs women. JAMA Oncol. 2019;5:1589–96.

[8] Muzaffar M, Kachare S, Vohra N. Impact of primary tumor surgery in stage IV male breast cancer. Clin Breast Cancer. 2017;17:e143–9.

[9] Dawood S, Broglio K, Gonzalez-Angulo AM, et al. Trends in survival over the past two decades among white and black patients with newly diagnosed stage IV breast cancer. J Clin Oncol. 2008;26:4891–8.

[10] Anderson WF, Althuis MD, Brinton LA, et al. Is male breast cancer similar or different than female breast cancer? Breast Cancer Res Treat. 2004;83:77–86.

[11] Press DJ, Miller ME, Liederbach E, et al. De novo metastasis in breast cancer: occurrence and overall survival stratified by molecular subtype. Clin Exp Metastasis. 2017;34:457–65.

[12] Ye L-J, Suo H-D, Liang C-Y, et al. Nomogram for predicting the risk of bone metastasis in breast cancer: a SEER population-based study. Transl Cancer Res. 2020;9:6710–9.

[13] Metzger-Filho O, Sun Z, Viale G, et al. Patterns of Recurrence and outcome according to breast cancer subtypes in lymph node-negative disease: results from international breast cancer study group trials VIII and IX. J Clin Oncol. 2013;31:3083–90.

[14] Xiong Y, Cao H, Zhang Y, et al. Nomogram-predicted survival of breast cancer brain metastasis: a SEER-based population study. World Neurosurg. 2019;128:e823–34.

[15] Duggan MA, Anderson WF, Altekruse S, et al. The Surveillance, Epidemiology, and End Results (SEER) program and pathology: toward strengthening the critical relationship. Am J Surg Pathol. 2016;40:e94–e102.

[16] Xie J, Ying YY, Xu B, et al. Metastasis pattern and prognosis of male breast cancer patients in US: a population-based study from SEER database. Ther Adv Med Oncol. 2019;11:1–12.

[17] Hagberg KW, Taylor A, Hernandez RK, et al. Incidence of bone metastases in breast cancer patients in the United Kingdom: results of a multi-database linkage study using the general practice research database. Cancer Epidemiol. 2013;37:240–6.

[18] Liu L, Chi YY, Wang AA, et al. Marital status and survival of patients with hormone receptor-positive male breast cancer: a Surveillance, Epidemiology, and End Results (SEER) population-based study. Med Sci Monit. 2018;24:3425–41.

[19] Wu Q, Li J, Zhu S, et al. Breast cancer subtypes predict the preferential site of distant metastases: a SEER based study. Oncotarget. 2017;8:27990–6.

[20] Bendell JC, Domchek SM, Burstein HJ, et al. Central nervous system metastases in women who receive trastuzumab-based therapy for metastatic breast carcinoma. Cancer. 2003;97:2972–7.

[21] Zhang M, Wang B, Liu N, et al. Nomogram for predicting preoperative regional lymph nodes metastasis in patients with metaplastic breast cancer: a SEER population-based study. BMC Cancer. 2021;21:565.

[22] Lu YJ, Yang Y, Yuan YH, et al. A novel nomogram based on SEER database for the prediction of liver metastasis in patients with small-cell lung cancer. Ann Palliat Med. 2020;9:3123–37.

[23] Purushotham A, Shamil E, Cariati M, et al. Age at diagnosis and distant metastasis in breast cancer-a surprising inverse relationship. Eur J Cancer. 2014;50:1697–705.

[24] James JJ, Evans AJ, Pinder SE, et al. Bone metastases from breast carcinoma: histopathological - radiological correlations and prognostic features. Br J Cancer. 2003;89:660–5.

[25] Lin Z, Yan S, Zhang J, Pan Q. A nomogram for distinction and potential prediction of liver metastasis in breast cancer patients. J Cancer. 2018;9:2098–106.

[26] Rudolph PM, Bonichon F, Frahm SO, et al. Prognostic significance of Ki-67 and topoisomerase II alpha expression in infiltrating ductal carcinoma of the breast. Breast Cancer Res Treat. 1999;1:61–71.

[27] Bozcuk H, Uslu G, Pestereli E, et al. Predictors of distant metastasis at presentation in breast cancer: a study also evaluating associations among common biological indicators. Breast Cancer Res Treat. 2001;68:239–48.