A Risk Stratification Model for Predicting Survival in Breast Cancer Patients With Lung Metastasis: A Study of the SEER Database and A Chinese Cohort

Xiang-Chen Han
Fudan University Shanghai Cancer Center
Yu-Chen Pei
Fudan University Shanghai Cancer Center
Yun-Jin Wang
Fudan University Shanghai Cancer Center
Xiao-Guang Li
Fudan University Shanghai Cancer Center
Xun-Xi Lu
Fudan University Shanghai Cancer Center
Lin-Wei Guo
Fudan University Shanghai Cancer Center
Bo-Yue Han
Fudan University Shanghai Cancer Center
Yue Gong
Fudan University Shanghai Cancer Center
Hong Ling
Fudan University Shanghai Cancer Center
Xin Hu (Xinhufd@163.com)
Fudan University Shanghai Cancer Center

Research Article

Keywords: breast cancer, lung metastases, nomogram, overall survival, breast cancer-specific survival, Chinese

Posted Date: March 10th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-255801/v1

Abstract

Background Distant metastases are the leading cause of death among breast cancer (BC) patients, and lung is the second preferential colonized sits. However, there has been no prognostic evaluation of BC patients with lung metastases so far. This research aimed to predict the overall survival (OS) and breast cancer-specific survival (BCSS) for BC patients with lung metastases and stratify them into different risk groups.

Result A total of 3888 patients from the SEER database were eligible for the subsequent analyses, with the training cohort of 1944 patients and the validation cohort I of 1944 patients. In addition, 374 patients from the Chinese XXX database were assigned as the validation II. Race, age, grade, subtype, number of metastatic sites, surgery, and chemotherapy were identified as determinant prognostic variables and integrated to construct the nomogram. Ultimately, a risk stratification model was established, and all patients were classified as three risk groups.

Conclusion We constructed prognostic prediction and risk stratification models for BC patients with lung metastases and validated it in both the SEER cohort and the Chinese cohort; these could assist clinicians in risk prediction, prognosis evaluation, and decision making on an individualized basis.

Introduction

Breast cancer (BC), accounting for 30% of the female cancers, is the most frequently diagnosed malignancy and second leading cause among women worldwide[1]. Despite the treatment breakthrough, its metastases at distant sites remain to be the primary cause of death among BC patients, reducing the 5-year survival rate from 98–27%[2]. Only approximately 6% of patients diagnosed as stage IV disease initially; while approximately 30% of patients diagnosed as early-stage disease initially but developed stage IV disease eventually[3]. Furthermore, owing to the increased detection of metastasis, the incidence rate of stage IV BC increased by 2.5% annually between 2001 to 2011[4].

It's well recognized that tumors, including breast cancer, tend to metastasize to certain organs[5]. In the case of BC, lung and bone are the most frequently colonized sites, and the liver and brain in a lower extend. Especially, BCs can be categorized into different molecular subtypes based on hormone receptors (HR) and human epidermal

growth factor receptor 2 (HER2) and growing evidence reveal that different molecular subtype prefers diverse sites of metastases[6–7]. To be specific, triple negative (HER2-/HR-), triple positive (HER2+/HR+) and HER2-enriched (HER2+/HR-) subtype preferentially metastasize to visceral organs including the lung, liver, and brain; while hormone receptor positive (HR+) subtype tends to metastasize to the bone[8–9].

Notably, among all metastatic BC patients, the lung is one of the most common colonized sits, with an autopsy incidence of 80%[10]. Usually, breast cancer lung metastasis (BCLM) patients elicit little or no

clinical symptoms until the metastatic tumors replace the vast of the lung. Despite varieties of treatments for BCLM patients, including endocrine therapy, chemotherapy, and target therapy, the prognostic outcomes of BCLM patients are still dismal, with the median OS time of 21 months[11]. Besides, BC is a heterogenous malignancy with diverse characteristics and prognosis. The prognosis endings of BC patients can be affected by the demographical and clinicopathological factors, including race, age, subtype, grade, metastatic sites, tumor-node-metastasis (TNM)[12–15]. For example, compared to bone metastases, lung, liver, and brain metastases exhibit worse prognosis[16]. According to a retrospective study, among BCLM patients, the survival outcomes of HR+/HER2- and HR+/HER2 + subtypes are better than those of HR-/HER2 + and HR-/HER2- subtype[11]. Thus, it's of great significance to evaluate the prognosis in BCLM patients.

In recent years, nomograms are widely used to conveniently and accurately predict the clinical outcomes of the particular population, especially among cancer patients[17–20]. Thus, nomograms could help the clinicians in prognosis evaluation and decision making in clinical practice, meeting our desire for personalized medicine. Compared with the conventional TNM staging system, nomograms could integrate all parameters related to survival to provide a more accurate estimation[21]. Recently, several nomograms have been constructed in metastatic BC patients[22–25]. Nevertheless, there is no nomogram built for BC patients with lung metastasis. In our research, we constructed and validated nomograms for those patients based on a large-scale population from the Surveillance, Epidemiology, and End Results (SEER) dataset.

Material And Methods

Study cohorts

Data for this study were obtained from the SEER database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2018 Sub (1975-2016 varying) and the latest Radiation/Chemotherapy Databases, Apr 2019 Updated (1975-2016). Since the SEER database didn't include detailed metastatic sites and molecular classification until 2010, we only enrolled patients diagnosed from 2010 to 2016. Moreover, we extracted 374 patients diagnosed from 2008 to 2018 in Chinese XXX database. The data screening was conducted based on the following inclusion and exclusion criteria (Figure1).

Inclusion criteria: (1) diagnosed between 2010 to 2016; (2) aged 18 years or older; (3) female patients; (4) diagnosed as malignant breast cancer by positive histology only; (5) histologic type: International Classification of Diseases for Oncology Third Edition(ICD-O-3) =8500-8543 (ductal, lobular and medullary carcinoma); (6) with certain follow-up data; (7) with lung metastasis.

Exclusion criteria: incomplete clinicopathologic information including race, grade, laterality, subtype, metastasis, surgery, and chemotherapy.

Statistical analyses

To construct, validate, and calibrate the nomogram, the eligible population were randomly divided into the training group and the validation group. Descriptive analysis was applied to depict the demographic, clinicopathological features, and socioeconomic features of the population. In the training cohort, OS was estimated in different subgroups via Kaplan-Meier plots and univariate prognostic parameters were compared by log-rank tests. Prognostic factors achieving significance at $P<0.05$ were enrolled for future multivariable analyses via the Cox proportional hazards regression model. The determinant prognostic variables identified by the multivariable analyses were chosen to generate the nomograms.

Considering the potential competitive risk factors, BCSS was also estimated via univariate and multivariate analyses in the training cohort. The probability of death was assessed by the cumulative incidence function (CIF) and the difference in CIF between subgroups was tested using the Find & Gray model.

Based on the outcomes of the univariate and multivariate analyses, nomograms were generated via the *rooms* and *survival* package to predict the 2-, 3- and 5-year probabilities of the OS and BCSS in patients with lung metastasis.

To validate and calibrate the nomograms, 1000 bootstrap resamples were conducted for internal validation within the training cohort and external validation with the validation cohorts. The Calibration curves were conducted to assess the ability of the nomogram to make unbiased estimates and the Harrell concordance index (C-index) was used to evaluate the discrimination performance via the *Hmisc* package.

Based on each patient's total risk score from the nomogram for OS prediction in the training cohort, a risk stratification model was established. All population were divided into three prognosis groups and the respective Kaplan-Meier plots were developed via the *survminer*, *ggplot2*, and *survival* package.

Two-sided $P<0.05$ were considered statistically significant. All statistical analyses were executed by SPSS Statistics 25.0 and R 4.0.2.

Results

Patient characteristics

After data selection, 3888 BC patients with lung metastasis were eligible for the subsequent analyses, with the training group of 1944 patients and the validation group I of 1944 patients. In addition, 374 patients from the Chinese XXX database were assigned as the validation II. The demographic and clinicopathological features of the population are shown in Table1. The distribution of each subgroup in the training cohort and the validation cohort I were similar enough, verifying that two cohorts were randomly divided. In the training cohort, 5.35% (104/1944), 37.14% (722/1944), 45.99% (894/1944), and 11.52% (224/1944) of the patients aged 18-39, 40-59, 60-79 and >80 respectively. Moreover, patients with

HER2-/HR-, HER2+/HR-, HER2-/HR+, and HER2+/HR+ subtype were accounted for 18.31% (356/1944), 10.24% (199/1944), 53.14% (1033/1944), 18.31% (356/1944), and only 692 patients (35.60%) had metastasis confined to the lung. The median survival time was 15 months.

For the validation cohort II, 19.52% (73/374), 58.02% (217/374), 20.86% (78/374), and 1.6% (6/374) of the patients aged 18-39, 40-59, 60-79 and >80, respectively. In addition, patients with HER2-/HR-, HER2+/HR-, HER2-/HR+, and HER2+/HR+ subtype were accounted for 21.66% (81/374), 17.38% (65/374), 43.05% (161/374), 17.91% (67/374), and only 207 patients (55.35%) had metastasis confined to the lung. The median survival time was 54 months.

Univariate and Multivariate analyses of Overall Survival

Outcomes from the univariate and multivariate analyses of the training cohort and the validation cohort are listed in Table2. The following prognostic parameters were identified determinant prognostic parameters for OS: age (40-59: HR 1.206, 95%CI 0.909-1.599; 60-79: HR 1.46, 95%CI 1.101-1.935; >80: HR 2.399, 95%CI 1.728-3.332; 18-39 as a reference), race (black, HR 1.291, 95%CI 1.109-1.503; others: HR 0.873, 95%CI 0.692-1.102; white as a reference), grade□III/IV: HR 1.696, 95%CI 1.478-1.948; I/II as a reference), subtype (HER2+/HR-: HR 0.413, 95%CI 0.325-0.524; HER2-/HR+: HR 0.390, 95%CI 0.328-0.464; HER2+/HR+: HR 0.267, 95%CI 0.216-0.331; HER2-/HR- as a reference), number of metastatic sites to liver, brain, and bone (1: HR 1.396, 95%CI 1.203-1.619; 2: HR 2.548, 95%CI 2.152-3.016; 3: HR 3.710, 95%CI 2.707-5.085; 0 as a reference), surgery (BCS: HR 0.552, 95%CI 0.436-0.699; mastectomy: HR 0.691, 95%CI 0.591-0.807; no as a reference), and chemotherapy(no/unknown: HR, 1.465, 95%CI 1.271-1.688, yes as a reference). Therefore, we selected these statistically significant prognostic parameters and established a nomogram to predict the 2-, 3-, and 5- year OS.

Breast cancer-specific survival

Assessments of the probabilities of breast cancer-specific death (BCSD) and non-breast cancer-specific death (NBCSD) among the training group are shown in Table3. Obviously, black, grade III/IV, HER2-/HR- subtype, 3 metastasis sites to the liver, bone, and brain, and not accepting surgery were only significantly correlated with higher two- and five- year cumulative incidence of BSCD ($P < 0.001$). Especially, 7 determinant prognostic factors for OS were also significantly correlated with the probabilities of BCSD ($P < 0.05$ for all outcomes), indicating that the previous outcome was not influenced by the competing events. All prognostic factors which were significantly associated with cumulative incidences of BCSD and statistically significant for OS via multivariate analysis simultaneously were selected for nomogram construction.

Construction, validation, and calibration of the nomograms

Nomograms integrated eight independent prognostic factors to predict the 2-, 3- and 5- year OS and BCSS in the training cohort (Figure2). Scores assigned for the factors in each subgroup are listed in Table4.Among all involved factors, 3 metastatic sites to the liver, brain, and bone had the highest score of

100, followed by subtype (HER2-/HR-: score 96), age (>80: score 69), the number of metastasis sites to the liver, brain, and bone (2: score 68), surgery (no: score 43), grade (III/UD: score 37), race (black: score 34), chemotherapy (no: score29). To predict the probability of 2-, 3-, and 5- year OS, we simply added up the scores of each variable to obtain the total score of individual patients and located the total score on the bottom scales.

The C-indexes for the nomograms predicting OS and BCSS in the training group were 0.707 (95%CI 0.690-0.724) and 0.698 (95%CI 0.702-0.732), suggesting acceptable discrimination performance. In the training group and the validation groups, the calibration curves of the nomograms predicting OS and BCSS (Figure3) fell on a 45° diagonal line, which indicated consistency between the nomogram prediction and recorded results for 2-year OS and 2-year BCSS in both the US SEER cohort and the Chinese XXX cohort.

Risk stratification model

Furthermore, we developed a risk stratification model according to the total score of individual patients from the nomogram and stratified them into three groups: low-risk group (791/1994, 39.67%; total score ≤150), intermediate-risk group (899/1994, 45.09%; total score 151-224), and high-risk group (284/1994, 14.24%; total score ≥225). In the training cohort, the median OS time in the low-, intermediate- and high-risk group was 46.1months, 28.7months, and 13.4months, respectively. In the validation cohort II, the median OS time in the low-, intermediate- and high-risk group was 83.5months, 50.9months, and 39.0months, respectively. The Kaplan-Meier survival plots of OS and BCSS for all cohorts, the training cohort, and the validation cohort I, and validation cohort II, achieved statistical significance ($P < 0.001$ for all outcomes), which illustrated the capacity of the risk stratification model to discriminate prognosis between the different risk subgroups (Figure4).

Discussion

Distant metastases are the primary cause of death among BC patients, and lung is the second preferential colonized sits for breast cancer[2]. Therefore, it is necessary to identify the factors related to prognosis and integrate them all to predict individual survival. Moreover, BC is a highly heterogeneous malignancy, and metastatic BC patients have multiple features different from other BC patients. Although numerous nomograms have been established for predicting the prognostic endings in the BC population, no specific nomogram has been reported for the population with lung metastasis yet[26–28]. In our work, we constructed and validated nomograms for predicting 2-, 3-, and 5-year OS and BCSS for BC patients with lung metastasis on the foundation of a large-scale population from the US SEER database and a Chinese cohort. Furthermore, a risk stratification model was developed according to the total scores of individual patients from the OS nomogram. The nomogram and risk stratification model could assist both clinicians and patients in personalized decision-making and clinical study design. For instance, high-risk patients need to accept additional systematic therapies, and the clinical follow-up should be narrowed to adjust the treatment protocols timely.

Although distant metastases are the primary cause of death among BC patients, some patients die from other causes. These competitive events might lead to the overestimation of the mortality risks. To assess the impact of competitive events, we introduced a competing-risk model[29]. The 2- and 5-year cumulative incidences of BCSD were 36.5% and 58.3%, respectively. Also, the 2- and 5-year probabilities of NBCSD were 12.3% and 20.4%, respectively, suggesting an approximately three-fold higher risk of BCSD than NBCSD. Notably, except for age and chemotherapy, all factors were found only significantly correlated with BCSD ($P < 0.001$ for all outcomes).

According to the analyses, we identified race, age, grade, subtype, surgery, chemotherapy, and number of metastatic sites as determinant prognostic parameters, which were in accordance with the previous publications[30–32]. Previous studies have highlighted that older, black, grade III/IV, and HR-/HER2- subtype probably have worse prognosis among stage IV BC patients, and our study confirmed that[11].

Besides, we found that patients undergoing surgery would probably have a better prognosis, prolonging both OS and BCSS time. Earlier studies reported that metastatic BC patients could benefit from initial breast surgery, as the surgery could lower the tumor burden, provide accurate pathological information, and alleviate clinical symptoms[33]. Moreover, a current meta-analysis integrated the results from 30 observation research and highlighted that locoregional surgery significantly prolonged OS time in metastatic BC patients, especially in patients with clear margins or limited disease burden[34]. Surprisingly, patients accepting BCS would probably have better prognosis than those accepting mastectomy, which might be explained that patients accepting BCS had lower tumor burden.

Surprisingly, the median OS time in the low-, intermediate- and high-risk group for the US SEER cohort and Chinese cohort was 46.1 and 83.5, 28.7 and 50.9 months, and 13.4 and 39.0months, respectively. The difference in the median OS time might have resulted from the small scale of the Chinese cohort, and large-scale, multicenter cohorts would be better. Among the Chinese validation cohort, the calibration curves of the nomograms fell on a 45° diagonal line and the Kaplan-Meier survival plots between different risk subgroups achieved statistical significance. Those indicated that the nomograms and risk stratification models generated according to the US SEER cohorts could also be utilized in the Chinese cohort to identify the high-risk population.

Nevertheless, some limitations should be taken into consideration. Firstly, a proportion of patients were excluded because of lacking certain follow-up data and complete clinicopathologic information for some important variables, such as grade and subtype, which may lead to some bias. Secondly, some prognostic parameters including Ki-67 positivity, multigene signature assessment, body mass index (BMI), Eastern Cooperative Oncology Group (ECOG) performance status and smoking status were not recorded in the SEER database, which might increase the robustness and effectiveness of the predictive model[36–37]. Thirdly, the detailed treatment was not available in the SEER database. Since the systematic treatments could provide better control of distant disease in BC, detailed chemotherapy, endocrine therapy, target therapy, and immunotherapy protocols are of great importance. Fourthly, the SEER database only records patients with distant metastasis. Therefore, patients with recurrence or later

metastasis during follow-up are not enrolled in this study. Last but not least, the retrospective study limited the application of our predictive model. Thus, further validation in a prospective cohort is needed before the nomograms are applied in clinical practice.

Conclusion

Using a large-scale population, we identified the independent prognostic parameters and constructed nomograms to predict the prognostic outcomes in BC patients with lung metastases in both the US and Chinses population. Our nomogram could visualize the assessment of the probability of 2-, 3-, and 5-year OS and BCSS. Furthermore, we established a risk stratification model to recognize the patients of high risk who need more personalized treatments.

Abbreviations

BC
breast cancer, ER = estrogen receptor, PR = progesterone receptor, HR = hormone receptor, HER-2 = human epidermal growth factor 2, OS = overall survival, BCSS = breast cancer-specific survival, BCLM = breast cancer lung metastasis, ICD-O-3 = International Classification of Disease for Oncology, Third Edition, SEER = Surveillance, Epidemiology, and End Results, IQR = interquartile range, CI = confidence interval, BSCD = breast cancer-specific death, NBCSD = non-breast cancer-specific death, US = United States.

Declarations

Ethics approval and consent to participate

This study did not involve animals.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

This retrospective study was approved by the Ethics Committee Review Board of Fudan University Shanghai Cancer Center (050432), and the need to obtain informed consent was waived.

Consent for publication

Not applicable.

Availability of data and materials:

Dataset for this study are available from the Surveillance, Epidemiology, and End Results (SEER) database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2018 Sub (1975-2016 varying).

Competing interests

The authors have declared that no conflicts of interest exist in this work.

Funding

This study was supported by a grant from the National Natural Science Foundation of China (No. 81803555 to X.G. Li).

Authors' contributions

Xin Hu, Ling Hong, and Xiang-chen Han designed this study. Xiang-chen Ham and Bo-yue Han collected data. Xiang-chen Han, Xiao-guang Li, Yu-chen Pei, Yun-jin Wang, Xun-xi Li, Lin-wei Guo, and Yue Gong analyzed the statistics and wrote the manuscript. Xin Hu and Ling Hong revised the manuscript. All authors have approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA A Cancer J Clin*. 2020;70(1):7–30. doi:10.3322/caac.21590
2. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA A Cancer J Clin*. 2019;69(6):438–451. doi:10.3322/caac.21583
3. Johnson RH, Chien FL, Bleyer A. Incidence of Breast Cancer With Distant Involvement Among Women in the United States, 1976 to 2009. *JAMA*. 2013;309(8):800. doi:10.1001/jama.2013.776
4. Fidler IJ. The pathogenesis of cancer metastasis: the "seed and soil" hypothesis revisited. *Nature Reviews Cancer*. 2003;3(6):453–458. doi:10.1038/nrc1098
5. DeSantis CE, Ma J, Jemal A. Trends in stage at diagnosis for young breast cancer patients in the United States. *Breast Cancer Res Treat*. 2019;173(3):743–747. doi:10.1007/s10549-018-5042-1
6. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406(6797):747–752. doi:10.1038/35021093
7. Smid M, Wang Y, Zhang Y, et al. Subtypes of Breast Cancer Show Preferential Site of Relapse. *Cancer Research*. 2008;68(9):3108–3114. doi:10.1158/0008-5472.CAN-07-5644
8. Kennecke H, Yerushalmi R, Woods R, et al. Metastatic Behavior of Breast Cancer Subtypes. *JCO*. 2010;28(20):3271–3277. doi:10.1200/JCO.2009.25.9820
9. Kast K, Link T, Friedrich K, et al. Impact of breast cancer subtypes and patterns of metastasis on outcome. *Breast Cancer Res Treat*. 2015;150(3):621–629. doi:10.1007/s10549-015-3341-3

10. Cummings MC, Simpson PT, Reid LE, et al. Metastatic progression of breast cancer: insights from 50 years of autopsies: Historical insights on metastatic breast cancer. *J Pathol*. 2014;232(1):23–31. doi:10.1002/path.4288
11. Xiao W, Zheng S, Liu P, et al. Risk factors and survival outcomes in patients with breast cancer and lung metastasis: a population-based study. *Cancer Med*. 2018;7(3):922–930. doi:10.1002/cam4.1370
12. Derks MGM, Bastiaannet E, van de Water W, et al. Impact of age on breast cancer mortality and competing causes of death at 10 years follow-up in the adjuvant TEAM trial. *European Journal of Cancer*. 2018;99:1–8. doi:10.1016/j.ejca.2018.04.009
13. Ren J-X, Gong Y, Ling H, Hu X, Shao Z-M. Racial/ethnic differences in the outcomes of patients with metastatic breast cancer: contributions of demographic, socioeconomic, tumor and metastatic characteristics. *Breast Cancer Res Treat*. 2019;173(1):225–237. doi:10.1007/s10549-018-4956-y
14. Bonafede MM, Miller JD, Pohlman SK, et al. Breast, Cervical, and Colorectal Cancer Screening: Patterns Among Women With Medicaid and Commercial Insurance. *American Journal of Preventive Medicine*. 2019;57(3):394–402. doi:10.1016/j.amepre.2019.04.010
15. Chu J, Yang D, Wang L, Xia J. Nomograms predicting survival for all four subtypes of breast cancer: a SEER-based population study. *Ann Transl Med*. 2020;8(8):544–544. doi:10.21037/atm-20-2808
16. Wang R, Zhu Y, Liu X, Liao X, He J, Niu L. The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer. *BMC Cancer*. 2019;19(1):1091. doi:10.1186/s12885-019-6311-z
17. Spiegel DY, Hong JC, Oyekunle T, et al. A Nomogram for Testosterone Recovery After Combined Androgen Deprivation and Radiation Therapy for Prostate Cancer. *International Journal of Radiation Oncology*Biology*Physics*. 2019;103(4):834–842. doi:10.1016/j.ijrobp.2018.11.007
18. Wang S, Yang L, Ci B, et al. Development and Validation of a Nomogram Prognostic Model for SCLC Patients. *Journal of Thoracic Oncology*. 2018;13(9):1338–1348. doi:10.1016/j.jtho.2018.05.037
19. Huang Y, Liang C, He L, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *JCO*. 2016;34(18):2157–2164. doi:10.1200/JCO.2015.65.9128
20. Callegaro D, Miceli R, Mariani L, Raut CP, Gronchi A. Soft tissue sarcoma nomograms and their incorporation into practice: Nomograms for Patients With Sarcoma. *Cancer*. 2017;123(15):2802–2820. doi:10.1002/cncr.30721
21. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *The Lancet Oncology*. 2015;16(4):e173-e180. doi:10.1016/S1470-2045(14)71116-7
22. Xiong Z, Deng G, Huang X, et al. Score for the Survival Probability in Metastasis Breast Cancer: A Nomogram-Based Risk Assessment Model. *Cancer Res Treat*. 2018;50(4):1260–1269. doi:10.4143/crt.2017.443
23. 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(12):2098–2106. doi:10.7150/jca.24445

24. Ouldamer L, Bendifallah S, Chas M, et al. Intrinsic and extrinsic flaws of the nomogram predicting bone-only metastasis in women with early breast cancer: An external validation study. *European Journal of Cancer*. 2016;69:102–109. doi:10.1016/j.ejca.2016.09.039
25. 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(11):2586–2594. doi:10.1002/cam4.1224
26. Pan X, Yang W, Chen Y, Tong L, Li C, Li H. Nomogram for predicting the overall survival of patients with inflammatory breast cancer: A SEER-based study. *The Breast*. 2019;47:56–61. doi:10.1016/j.breast.2019.05.015
27. Diao J, Ma L, Sun M, et al. Construction and validation of a nomogram to predict overall survival in patients with inflammatory breast cancer. *Cancer Med*. 2019;8(12):5600–5608. doi:10.1002/cam4.2470
28. Xie X, Tan W, Chen B, et al. Preoperative prediction nomogram based on primary tumor miRNAs signature and clinical-related features for axillary lymph node metastasis in early-stage invasive breast cancer: A nomogram for breast cancer ALN metastasis prediction. *Int J Cancer*. 2018;142(9):1901–1910. doi:10.1002/ijc.31208
29. Southern DA, Faris PD, Brant R, et al. Kaplan–Meier methods yielded misleading results in competing risk scenarios. *Journal of Clinical Epidemiology*. 2006;59(10):1110–1114. doi:10.1016/j.jclinepi.2006.07.002
30. Kim Y-J, Kim J-S, Kim IA. Molecular subtype predicts incidence and prognosis of brain metastasis from breast cancer in SEER database. *J Cancer Res Clin Oncol*. 2018;144(9):1803–1816. doi:10.1007/s00432-018-2697-2
31. Li X, Yang J, Peng L, et al. Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast Cancer Res Treat*. 2017;161(2):279–287. doi:10.1007/s10549-016-4059-6
32. Dreyer MS, Nattinger AB, McGinley EL, Pezzin LE. Socioeconomic status and breast cancer treatment. *Breast Cancer Res Treat*. 2018;167(1):1–8. doi:10.1007/s10549-017-4490-3
33. Thomas A, Khan SA, Chrischilles EA, Schroeder MC. Initial Surgery and Survival in Stage IV Breast Cancer in the United States, 1988–2011. *JAMA Surg*. 2016;151(5):424. doi:10.1001/jamasurg.2015.4539
34. Xiao W, Zou Y, Zheng S, et al. Primary tumor resection in stage IV breast cancer: A systematic review and meta-analysis. *European Journal of Surgical Oncology*. 2018;44(10):1504–1512. doi:10.1016/j.ejso.2018.08.002
35. Street W. Cancer Facts & Figs. 2019. Published online 1930:76.
36. Li H, Sun X, Miller E, et al. BMI, reproductive factors, and breast cancer molecular subtypes: A case-control study and meta-analysis. *Journal of Epidemiology*. 2017;27(4):143–151. doi:10.1016/j.je.2016.05.002

37. Chen X, He C, Han D, et al. The predictive value of Ki-67 before neoadjuvant chemotherapy for breast cancer: a systematic review and meta-analysis. *Future Oncology*. 2017;13(9):843–857. doi:10.2217/fon-2016-0420

Tables

Table1. Demographic and clinical characteristics of the cohort

Demographic and clinical characteristic	Training cohort		Validation cohort □		Validation cohort □	
	N=1944		N=1944		N=374	
Age at diagnosis						
18-39	104	(5.35%)	114	(5.86%)	73	(19.52%)
40-59	722	(37.14%)	708	(36.42%)	217	(58.02%)
60-79	894	(45.99%)	880	(45.27%)	78	(20.86%)
>80	224	(11.52%)	242	(12.45%)	6	(1.60%)
Race						
White	1415	(72.79%)	1409	(72.48%)	0	(0.00%)
Black	366	(18.83%)	344	(17.70%)	0	(0.00%)
Others†	163	(8.38%)	191	(9.83%)	374	(100.00%)
Laterality						
Left	983	(50.57%)	992	(51.03%)	182	(48.66%)
Right	961	(49.43%)	952	(48.97%)	192	(51.34%)
Histology						
IDC	1724	(88.68%)	1710	(87.96%)	334	(89.30%)
ILC	137	(7.05%)	149	(7.66%)	27	(7.22%)
Others‡	83	(4.27%)	85	(4.37%)	13	(3.48%)
Grade§						
□/□	873	(44.91%)	878	(45.16%)	181	(48.40%)
□/UD	1071	(55.09%)	1066	(54.84%)	193	(51.60%)
Subtype						
HER2-/HR-	356	(18.31%)	329	(16.92%)	81	(21.66%)
HER2+/HR-	199	(10.24%)	198	(10.19%)	65	(17.38%)
HER2-/HR+	1033	(53.14%)	1063	(54.68%)	161	(43.05%)
HER2+/HR+	356	(18.31%)	354	(18.21%)	67	(17.91%)
Tumor size (cm)						
<2	221	(11.37%)	208	(10.70%)	130	(34.76%)

2-5	665	(34.21%)	714	(36.73%)	204	(54.55%)
>5	844	(43.42%)	812	(41.77%)	26	(6.95%)
Unknown	214	(11.01%)	210	(10.80%)	14	(3.74%)
Lymph node infiltration						
Negative	408	(20.99%)	425	(21.86%)	126	(33.69%)
Positive	1419	(72.99%)	1406	(72.33%)	244	(65.24%)
Unknown	117	(6.02%)	113	(5.81%)	4	(1.07%)
Number of metastatic sites to liver, bone and brain						
0	692	(35.60%)	685	(35.24%)	207	(55.35%)
1	802	(41.26%)	824	(42.39%)	111	(29.68%)
2	381	(19.60%)	378	(19.44%)	48	(12.83%)
3	69	(3.55%)	57	(2.93%)	8	(2.14%)
Surgery						
No	1387	(71.35%)	1406	(72.33%)	0	(0.00%)
BCS	157	(8.08%)	131	(6.74%)	39	(10.43%)
Mastectomy	400	(20.58%)	407	(20.94%)	407	(108.82%)
Radiation						
Yes	565	(29.06%)	557	(28.65%)	202	(54.01%)
No	1379	(70.94%)	1387	(71.35%)	172	(45.99%)
Chemotherapy						
Yes	1213	(62.40%)	1204	(61.93%)	334	(89.30%)
No	731	(37.60%)	740	(38.07%)	40	(10.70%)
Survival months						
Median (IQR)	15	(6-30)	15	(6-31)	54	(36-77)

Abbreviations: IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; IQR, interquartile range.

Others†, including American Indian/AK Native and Asian/Pacific Islander.

Others‡, including medullary carcinoma.

Grade§, histopathological grading. UD= undifferentiated.

Table2. Univariate and Multivariate analyses of OS in the Training Group						
Variable	Univariate analysis	Multivariate analysis				
	P Value	HR	95% CI			P Value
Age at diagnosis	<0.001***					<0.001***
18-39		Reference				
40-59		1.206	0.909	-	1.599	0.193
60-79		1.46	1.101	-	1.935	0.009
>80		2.399	1.728	-	3.332	<0.001***
Race	<0.001***					0.001
White		Reference				
Black		1.291	1.109	-	1.503	0.001**
Others		0.873	0.692	-	1.102	0.254
Laterality	0.209					
Left						
Right						
Histology	0.45					
IDC						
ILC						
Others						
Grade	<0.001***					<0.001***
□/□		Reference				
□/□		1.696	1.478	-	1.948	
Subtype	<0.001***					<0.001***
HER2-/HR-		Reference				
HER2+/HR-		0.413	0.325	-	0.524	<0.001***
HER2-/HR+		0.39	0.328	-	0.464	<0.001***
HER2+/HR+		0.267	0.216	-	0.331	<0.001***
Tumor size (cm)	0.2					
<2						

2-5						
>5						
Unknown						
Lymph node infiltration	0.4					
Yes						
No						
Unknown						
Number of metastatic sites to liver, bone and brain	<0.001***					<0.001***
0		Reference				
1		1.396	1.203	-	1.619	<0.001***
2		2.548	2.152	-	3.016	<0.001***
3		3.71	2.707	-	5.085	<0.001***
Surgery	<0.001***					<0.001***
No		Reference				
BCS		0.552	0.436	-	0.699	<0.001***
Mastectomy		0.691	0.591	-	0.807	<0.001***
Radiation	0.8					
Yes						
No/Unknown						
Chemotherapy	<0.001***					<0.001***
Yes		Reference				
No/Unknown		1.465	1.271	-	1.688	

*, two-sided p values smaller than 0.05; **, two-sided p values smaller than 0.01; ***, two-sided p values smaller than 0.001.

Table3. Two- and Five-year Cumulative Incidences of Death in the Training Group

Variables	Cumulative incidence of breast cancer-specific death			Cumulative incidence of non-breast cancer-specific death		
	2y	5y	P Value	2y	5y	P Value
Age at diagnosis			<0.001***			<0.001***
18-39	0.383	0.710		0.069	0.119	
40-59	0.426	0.698		0.026	0.064	
60-79	0.437	0.652		0.066	0.112	
>80	0.501	0.726		0.139	0.198	
Race			<0.001***			0.496
White	0.575	0.720		0.069	0.107	
Black	0.674	0.802		0.081	0.121	
Others	0.504	0.653		0.058	0.163	
Grade			<0.001***			0.176
□/□	0.333	0.590		0.066	0.125	
□/□	0.523	0.758		0.054	0.087	
Subtype			<0.001***			0.026*
HER2-/HR-	0.739	0.886		0.055	0.060	
HER2+/HR-	0.416	0.590		0.066	0.078	
HER2+/HR-	0.387	0.655		0.066	0.134	
HER2+/HR+	0.303	0.583		0.041	0.051	
Number of metastatic sites to liver, brain, and bone			<0.001***			0.389
0	0.347	0.601		0.074	0.118	
1	0.415	0.671		0.054	0.108	
2	0.591	0.814		0.045	0.076	
3	0.715	0.759		0.064	0.064	
Surgery			<0.001***			0.206
No	0.471	0.716		0.064	0.113	

BCS	0.348	0.561		0.056	0.107	
Mastectomy	0.360	0.635		0.045	0.086	
Chemotherapy			0.025*			<0.001***
Yes	0.424	0.686		0.037	0.082	
No/Unknown	0.459	0.673		0.094	0.140	

*, two-sided p values smaller than 0.05; **, two-sided p values smaller than 0.01; ***, two-sided p values smaller than 0.001.

Table4. Score assignment for overall survival (OS) and breast cancer-specific survival (BCSS)		
Variables and Scores	Score for OS	Score for BCSS
Age at diagnosis		
18-39	0	0
40-59	17	23
60-79	31	31
>80	69	64
Race		
White	10	11
Black	34	32
Others	0	0
Grade		
□/□	0	0
□/□	37	37
Subtype		
HER2-/HR-	96	94
HER2+/HR-	28	23
HER2+/HR-	25	22
HER2+/HR+	0	0
Number of metastatic sites to liver, brain, and bone		
0	0	0
1	24	26
2	68	70
3	100	100
Surgery		
No	43	42
BCS	0	0
Mastectomy	16	18
Chemotherapy		

Yes	0	0
No/Unknown	29	25

Figures

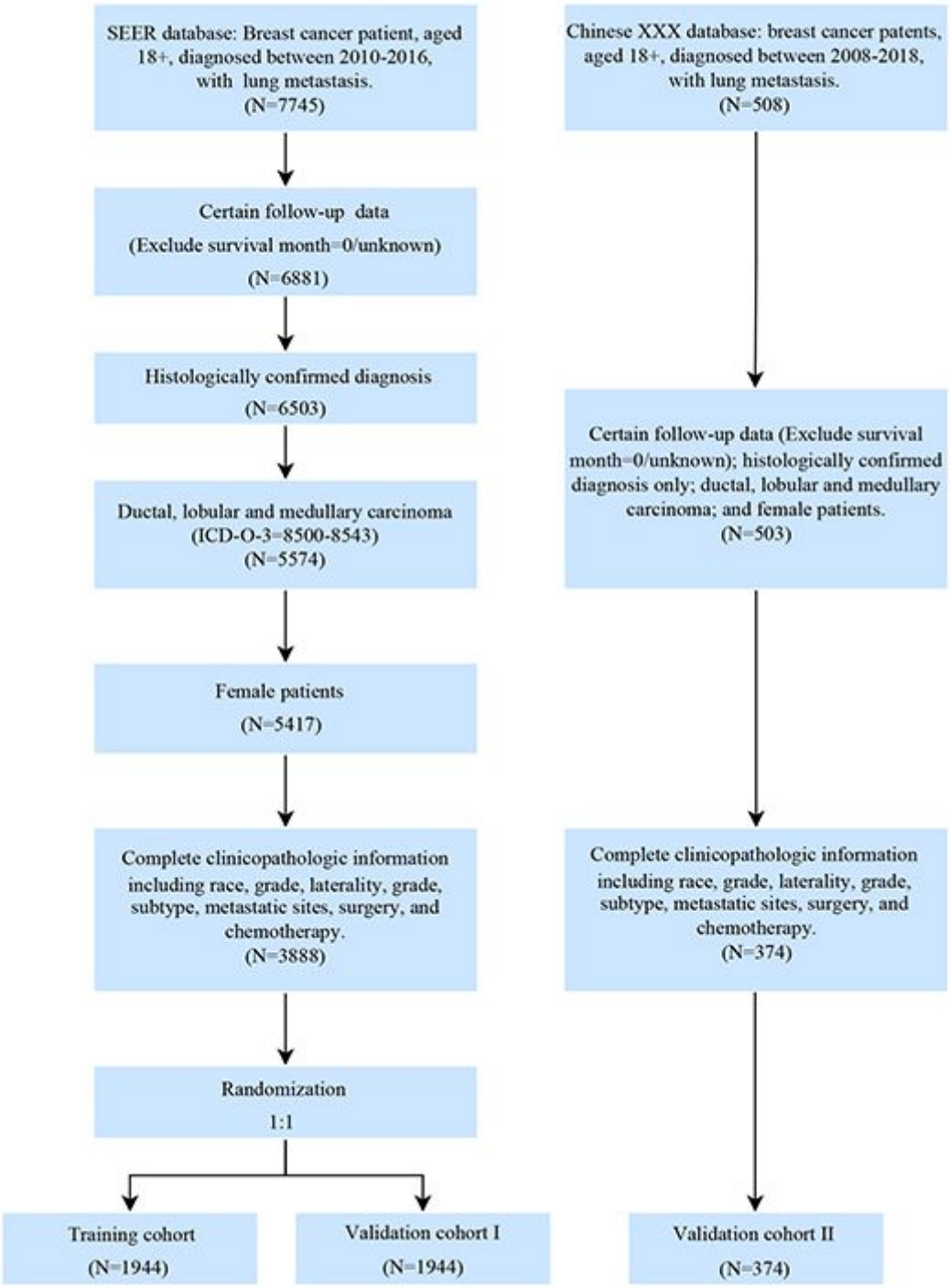

Figure 1

The flowchart for screening eligible patients. Abbreviations: ICD-O-3, International Classification of Disease for Oncology, Third Edition.

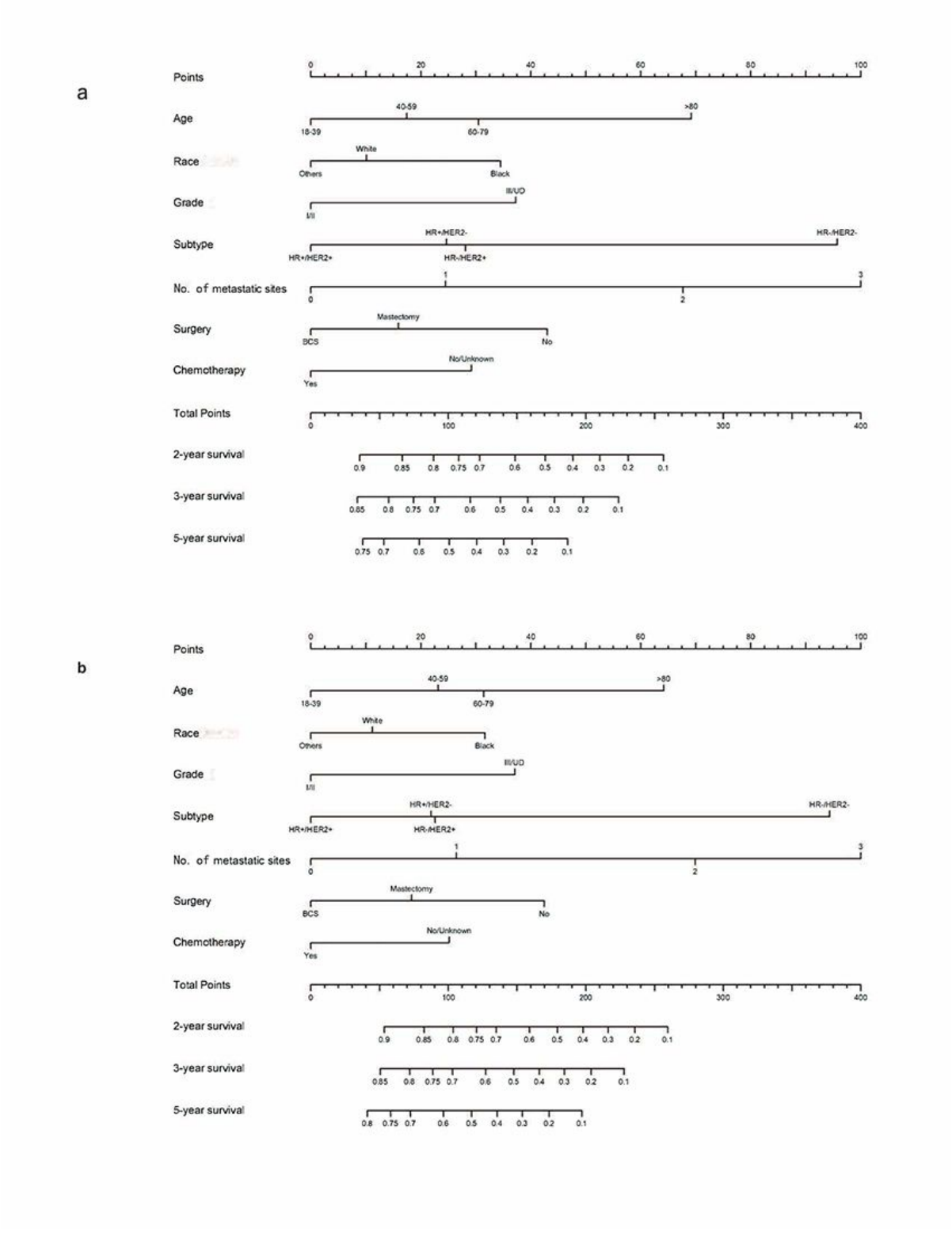

Figure 2

Nomograms predicting 2-, 3-, 5-year overall survival (a) and breast cancer-specific survival (b) in the training cohort.

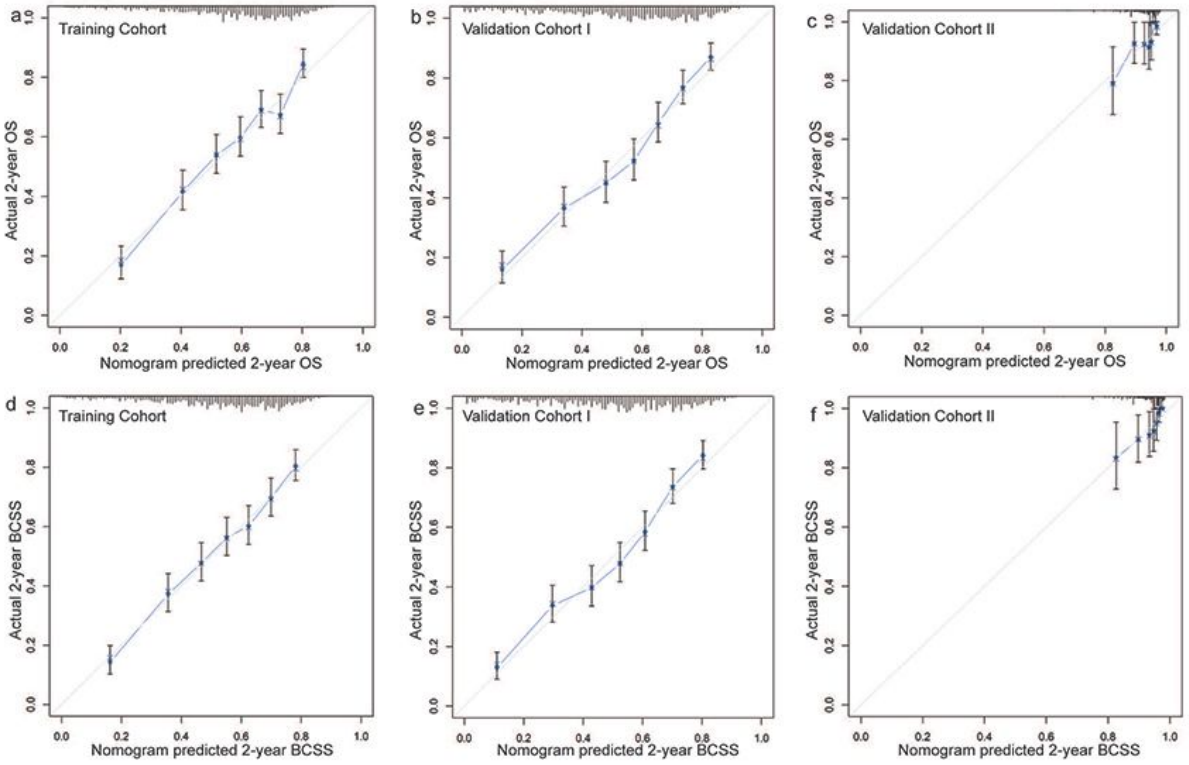

Figure 3

Calibration diagrams for the 2-year overall survival (OS) and breast cancer-specific survival (BCSS) in the training cohort (a, d), the validation cohort I (b, e), and the validation cohort II (c, f). The calibration curves fell on a 45° diagonal line. The perpendicular line indicates 95% confidence intervals.

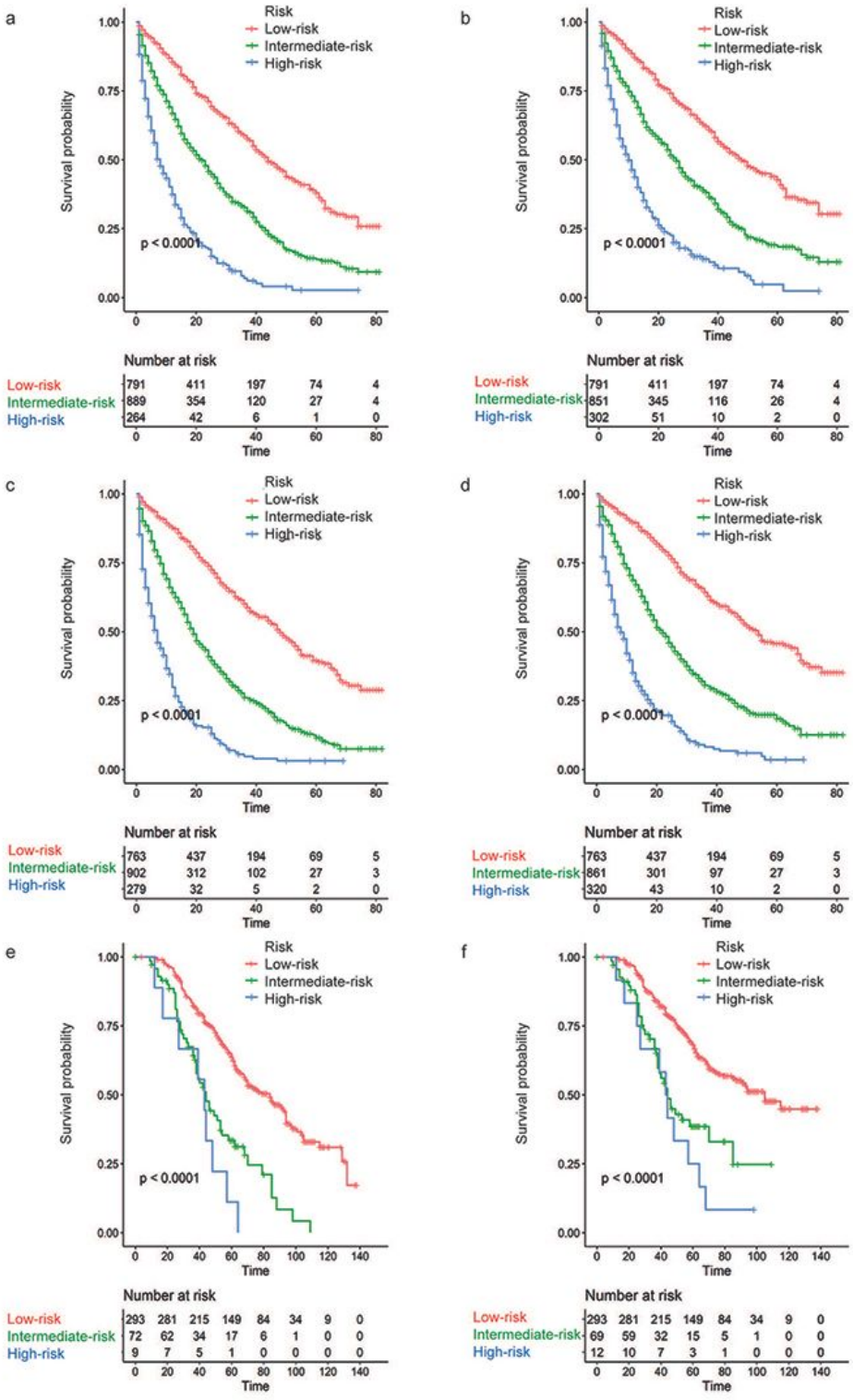

Figure 4

Kaplan-Meier curves of low-, intermediate- and high-risk groups for overall survival (OS) and breast cancer-specific survival (BCSS) in the training cohort (a, b), the validation cohort I (c, d), and the validation cohort II (e, f).