A Nomogram Predicting Survival in Patients with Breast Ductal Carcinoma in Situ with Microinvasion

Yi-Zi Zheng (14111230029@fudan.edu.cn)
Shenzhen Second People's Hospital https://orcid.org/0000-0002-3847-8494
Hong-Bin Qin
The People's Hospital of Hechi
Zi-Zheng Li
The People's Hospital of Hechi
He-Sheng Jiang
Fudan University Shanghai Cancer Center
Greg Zhang
The University of Texas Health Science Center at Houston
Shi-Wei Yang
Shenzhen Second People's Hospital
Xian-Ming Wang
Shenzhen Second People's Hospital
Yang-Chun Xu
The People's Hospital of Hechi
Zhen-Han Deng
Shenzhen Second People's Hospital
Guo-Wen Liu
Shenzhen Second People's Hospital

Research

Keywords: breast cancer, ductal carcinoma in situ, microinvasion, nomogram

Posted Date: February 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-218104/v1

Abstract

Background: Ductal carcinoma in situ with microinvasion (DCISM) can be challenging to balance the risks of overtreatment versus undertreatment. We aim to identify prognostic factors in patients with DCISM and construct a nomogram to predict breast cancer-specific survival (BCSS).

Methods: Women diagnosed with DCISM were selected from the Surveillance, Epidemiology and End Results database (1998-2015). Clinical variables and tumor characteristics were evaluated and Cox proportional-hazards regression model was performed. A nomogram was constructed from the multivariate logistic regression model to combine all the prognostic factors to predict the prognosis of DCISM patients at 5 years, 10 years, and 15 years.

Results: We identified 5,438 total eligible breast cancer patients with a median and max survival time of 78 and 227 months, respectively. Here, patients with poorer survival outcomes were those diagnosed between 1988-2001, African-American race, under 40 years of age, higher tumor N stage, progesterone receptor-negative tumor, and received no surgery (all $P < 0.05$). The nomogram was constructed by the seven variables and passed the calibration and validation steps. The area under the receiver operating characteristic (ROC) curve (AUC) of both the training set and the validating set (5-year AUC: 0.77 and 0.88, 10-year AUC: 0.75 and 0.73, 15-year AUC: 0.72 and 0.65) demonstrated excellent reliability and robust performance.

Conclusion: Our current study is the first to construct nomograms of patients with DCISM which could help physicians identify breast cancer patients that more likely to benefit from more intensive treatment and follow-up.

Background

Ductal carcinoma in situ (DCIS) with microinvasion (DCISM) is a mostly preinvasive breast carcinoma with a small component of invasive disease (presence of one or more foci of stromal invasion, none exceeding 1 mm in size) and presumably has a low but plausible risk of metastasis[1, 2]. Tumors with any invasive foci of 1mm or larger in size are defined as invasive carcinoma[1, 2]. Microinvasive carcinoma is an uncommon disease, accounting for a mere 1% of all breast cancer diagnoses[3-5]; furthermore, tumor microinvasion is found in association with only approximately 5–10% of DCIS cases[6-8]. Microinvasive cancer is rarely ever seen in the absence of an adjacent in situ lesion[6]. This may be due to difficulty visualizing an isolated 1-mm invasive component, whereas an adjacent in situ lesion will dramatically enhance its detectability. Consequently, microinvasive carcinoma is usually described as “DCIS with microinvasion” despite the presence of DCIS not being necessary. Although DCISM patients account for only a small proportion of total breast cancer cases, the incidence of DCISM continues to increase along with a very significant rise in DCIS as a result of increased detection of breast cancer with the widespread adoption of mammography screening[9, 10].

Current guidelines from the National Comprehensive Cancer Network (NCCN) recommend DCIS treatment and systemic therapy utilization for the majority of DCISM cases, which more closely reflects the therapeutic guidelines for DCIS than for that of invasive carcinoma[11]. However, several years ago it was recommended that patients with microinvasive carcinoma be treated the same as patients with small invasive cancers[12]. While surgery is the standard treatment in DCIS and the majority of invasive carcinomas, additional treatment options vary quite widely between the two entities. Most notably, adjuvant chemotherapy is part of the national treatment guidelines for many invasive breast but is not recommended for DCIS[13]. Given that DCISM is relatively rare

compared to pure DCIS and most invasive ductal carcinomas, there exists limited and controversial data regarding its tumor biology and diseases prognosis that serves to guide disease management and patient counseling. Several single-institution retrospective studies have reported clinical features, management, and prognostic implications for DCISM, but yield conflicting results[14-16]. Although DCIS, DCISM, and T1a invasive ductal carcinoma all have generally excellent prognosis, some population-based studies have revealed that DCISM more closely resembles small invasive carcinoma than pure DCIS and many practitioners are treating it accordingly as such[17, 18]. Breast cancer, even with microinvasion, is a very heterogeneous disease characterized by diverse histopathologic and molecular features that are associated with distinct clinical outcomes. As a result, it can be challenging to balance the potential risks of overtreatment versus undertreatment in DCISM.

The American Joint Committee for Cancer (AJCC) staging system is a widely used tool for clinicians to predict disease outcomes and guide therapeutic decision making[19, 20]. However, given the many variables that influence the course of cancer, a prognosis based on the AJCC staging system alone is simply insufficient. A precise estimate of DCISM mortality is required to evaluate the clinical implications of this early-stage cancer and guide individualized therapeutic approaches. Nomograms, with the ability to generate an individual probability of a clinical event by integrating biological and clinical variables, help fulfill this requirement and aid in the development of personalized medicine[21-23]. There are currently no studies constructing a nomogram for DCISM female breast cancer. To address this issue, this study aims to establish a comprehensive and reliable prognostic model of DCISM by building a nomogram to better understand the risk factors and prognosis. To obtain a sufficient number of DCISM cases, the Surveillance, Epidemiology and End Results (SEER) cancer database of the National Cancer Institute was used in this study.

Materials And Methods

Source of Data

Study data was obtained from the SEER database of the National Cancer Institute, an open access resource for epidemiologic and survival analyses of various cancers, consisting of a collection of 18 high quality population-based cancer registries with very high estimated completeness of reporting. All data is publicly available and de-identified, and therefore exempted from the review of an Institutional Review Board. SEER database data do not require informed consent.

The SEER*Stat software from the National Cancer Institute (Surveillance Research Program, National Cancer Institute SEER*Stat software, http://www.seer.cancer.gov/seerstat) (Version 8.1.5) was used to identify eligible patients with the following inclusion criteria: female, diagnosed between 1988-2015, pathological diagnosis of breast ductal carcinoma, unilateral breast cancer, stage T1mic, one primary site only, and known age at diagnosis. Information regarding the human epidermal growth factor receptor-2 (HER2/neu) status is only available in the SEER database from 2010 onwards; therefore, HER2 variable was not included in the analysis. Patients diagnosed with breast cancer after 2015 were excluded to ensure adequate follow-up time. The pathological diagnosis was based on the primary site and according to the International Classification of Disease for Oncology, Third Edition (ICD-O-3). Breast cancer-specific survival (BCSS) was the primary study outcome of the SEER data, which was calculated as the time period from the date of diagnosis to the date of breast cancer-

specific death. The causes of death were categorized as either breast cancer related or non-breast cancer related. Patients who died of non–breast cancer related causes were censored regarding the date of death.

Nomogram Development

The following clinical variables were extracted for the study: year of diagnosis, age, marital status, race, N stage (derived from AJCC stage group 6th edition), primary site, laterality, grade, estrogen receptor (ER) status, progesterone receptor (PR) status, surgery, chemotherapy, radiation. Continuous predictors were tested for linearity and converted to categorical variables if the relationship was determined to be nonlinear. Categorical variables were collapsed over categories, with no significant differences. For nomogram construction and validation, all cases were randomly divided into training (n = 3,806) and validating (n = 1,632) cohorts with a ratio of 7:3[24]. Univariate and multivariate Cox regression were then used to screen for variables that significantly correlated with BCSS in the training group. After backwards stepdown validation, predictors that remained in the model (*P* value less than 0.05) were year of diagnosis, age, race, N stage, PR status, surgery, and chemotherapy. The resulting multivariate Cox regression model was used to calculate risk score and build the final nomogram prognostic model.

Model Validation

The validity of the nomogram was tested by discrimination and calibration[21]. The discrimination was estimated by the area under the receiver operating characteristic (ROC) curve (AUC) [25]. The theoretical value of the AUC is between 0 and 1; an AUC larger than 0.5 indicates prediction performance better than random chance. Calibration curves were plotted to evaluate the consistency between predicted and actual survival rates at 5, 10, and 15 years [22]. A perfect prediction would result in a 45-degree calibration curve (i.e., the identity line).

Other Statistical Methodologies

To account for differences in baseline characteristics across the groups, we matched each patient who received chemotherapy to another patient who did not using the following predetermined factors: year of diagnosis, age, marital status, race, N stage, primary site, laterality, grade, ER status, PR status, surgery, chemotherapy, radiation. Propensity score matching method was utilized and the matching quality was tested. Kaplan–Meier curves, with the corresponding results of log-rank tests, were constructed for breast cancer-specific survival. The same methodology was carried out for patients receiving radiation therapy. All statistical analyses were performed in SPSS (version 24.0; IBM Corp, Armonk, NY, USA) or R environment (version 3.4.0; Vienna, Austria; http://www.R-project.org). All tests were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Clinicopathological Characteristics of Patients

Application of the aforementioned inclusion and exclusion criteria resulted in a final study population of 5,438 DCISM cases (Figure 1). These cases were randomly divided into two distinct groups: 3,806 cases were used as the training cohort, while 1,632 cases were used as the validating cohort. The follow-up time ranged from 0 to 227 months (median 78 months) for the training cohort and from 0 to 226 months (median 78 months) for the testing cohort. Patient, disease, and treatment characteristics for the study population are summarized in Table 1. The demographic and clinical variables were similar in the training and validating groups. The majority of patients

were diagnosed between 2002-2015, over 40 years of age, Caucasian, tumor grade II-III, ER positive, N0-N1 stage, and had undergone surgery and chemotherapy.

Building Nomogram Prognostic Model in Training Cohort

In the univariate analysis, each of the following variables significantly increased the BCSS: “diagnosed in 2002-2015”, “age between 40 and 70”, “married”, “Caucasian”, “N0 stage”, “grade I and II”, “PR positive”, “received surgery”, “no chemotherapy” and “received radiotherapy” (Table 2). After stepwise selection via multivariate analysis to further remove potential redundancies, the year of diagnosis, age, race, N stage, PR status, surgery, and chemotherapy were used in the final nomogram model (coefficients summarized in Table 2). The final risk scores for 5-year, 10-year, and 15-year BCSS were calculated by adding up the score of each item using the nomogram depicted in Figure 2. It was demonstrated that surgery contributed the most to prognosis, followed by N stage, age, race, chemotherapy, year of diagnosis, and lastly PR status.

Validation and Calibration of the Nomogram

The proposed nomogram was finally validated by discrimination and calibration measures in the independent testing set. The receiver operating characteristic (ROC) curves were plotted both internally and externally in the training and validating sets (Figure 3A and 3B). In the training set, the AUC for 5-year, 10-year, and 15-year BCSS were 0.77, 0.75 and 0.72, respectively. In the validating set, the AUC for 5-year, 10-year, and 15-year BCSS were 0.88, 0.73 and 0.65, respectively. This confirms the relatively strong prognostic power of the proposed nomogram. A calibration curve at 5 years (Figure 3C), 10 years (Figure 3D) or 15 years (Figure 3E) also showed high consistency between predicted probability and actual proportion of BCSS. The bias-corrected curve as well as the apparent curve were close to the ideal curve which falls along the 45-degree line, demonstrating the robustness of this nomogram.

Statistical Matching for Chemotherapy and Radiotherapy

Chemotherapy and radiotherapy were both commonly applied adjuvant therapies for treatment of breast cancer. Therefore, survival analyses were additionally performed for these two important variables. To ensure that differences in outcome were not attributed to baseline differences in demographic and clinical characteristics across the therapeutic groups, we performed a 1:1 (chemotherapy: no chemotherapy) matched case-control analysis using the propensity score-matching method. We obtained a group of 726 patients with 363 patients from each chemotherapy group (Figure 4A). Here, we found that chemotherapy was associated with a better prognosis of DCISM (Figure 4A, $P = 0.018$). The same analysis was performed for radiotherapy with a group of 1,588 patients with 794 patients in each radiotherapy group (Figure 4B). From this, we determined that radiotherapy was not associated with BCSS of DCISM (Figure 4B, $P = 0.872$).

Discussion

Because DCISM constitutes a small minority of cases of breast cancer, it has been difficult to definitively characterize its biological behavior, prognostic factors, and outcomes of multimodality therapy among patients. Previous studies have reported the prognostic implications and clinical management for DCISM, but the therapeutic recommendations proposed in microinvasive breast carcinomas are highly varied and remain controversial[14-16, 26]. Recent medical literature shows that current treatment patterns and prognosis of DCISM

are comparable to those with small volume invasive ductal carcinoma[17, 18]. DCISM breast cancer is a quite heterogeneous disease and could be associated with distinct clinical outcomes. It remains challenging to find a proper, balanced treatment. In this study, a nomogram prognostic model was developed and validated using a large cohort of breast DCISM cases across the United States. Based on routinely available demographic, staging, and treatment information, this nomogram predicts the survival probability for individual DCISM patients and contributes to the development of personalized medicine.

In our present study, we constructed a comprehensive model based on a combination of various risk factors to predict prognosis of breast DCISM. The seven variables include age, race, year of diagnosis, AJCC N stage, PR status, chemotherapy, and surgery were kept in this nomogram after multivariate Cox regression screening and backward stepwise selection; these were all readily available information in the clinical database. Measured by the concordance index, the nomogram passed the discrimination step with an AUC of 0.77, 0.75 and 0.72 (for 5-, 10-, 15-year BCSS, respectively) in the training set and 0.88, 0.73 and 0.65 (for 5-, 10-, 15-year BCSS, respectively) in the validating set, suggesting a decent capability of discerning the breast cancer-specific death event most of the time. Consider a 50-year-old African-American female patient with PR-negative, N1 stage DCISM receiving lumpectomy and chemotherapy in the year 2010. This patient's nomogram calculates that the 10-year probability of BCSS (AUC of 0.73) is 87% (Figure 2; nomogram calculations are as follows: age = 2, corresponding to score of 0; race =2, corresponding to score of 35; year of diagnosis = 2, corresponding to score of 0; AJCC N stage = 1, corresponding to score of 42.5; surgery = 1, corresponding to score of 0; chemotherapy = 1, corresponding to score of 32; PR status = 0, corresponding to score of 22; the sum equals 131.5 total points, corresponding to a 10-year BCSS of 87% [breast cancer-specific death of 13%], and 15-year BCSS of 82% [breast cancer-specific death of 18%]). At the 10-year time point, this patient has a 13% risk of breast cancer-specific death, calculated by a nomogram that can identify a breast cancer-specific death event 73% of the time. As depicted by the calibration plots in Figure 3, the apparent and bias-corrected curve both fit well with the ideal diagonal line. Additionally, taking 15-year BCSS as an example, this nomogram is more accurate at predicting BCSS of 90% than 97% (Figure 3E. On the calibration plot for the 15-year outcome, at a BCSS of 90%, the red circle overlaps with the blue dotted line indicating near-perfect calibration; however, at a BCSS of 97%, the red circle and blue dotted line do not overlap. As characterized by the confidence intervals in calibration plots, there obviously lies an additional degree of uncertainty in a nomogram estimation. Thus, we concluded that this nomogram model is nevertheless quite reliable and robust in making accurate assessments and predictions.

The prognostic factors described in this study were basically consistent with findings of previous studies. Younger age, lymph node metastasis, multifocality, positive hormone receptor status have all previously been shown to be of significant relevance to the prognosis of DCISM patients[18, 26-28]. Excluding the therapeutic factors, AJCC 6th edition N stage contributes the most to the final risk score (Figure 2), with clear distinctions between two adjacent N stages. The significant contribution of N stage to this nomogram strongly suggests that certain subsets of breast cancer may have an enhanced propensity to metastasize, exhibiting a worse prognosis even when the primary lesion is very small[29-31]. It is evident from Figure 2 that patients with DCISM treated by lumpectomy had the same or even slightly lower risk scores than those who were treated by mastectomy. Data looking specifically at DCISM is quite limited. The National Surgical Adjuvant Breast and Bowel Project (NSABP) B-06 has shown that stage I and II breast cancer patients who underwent lumpectomy with subsequent radiation had the same rate of survival as those who underwent mastectomies. A recent study based on well-matched, contemporary data revealed that breast-conserving therapy was associated with superior overall survival compared to mastectomy for early-stage breast cancer[32], and is consistent with the results in our current study.

Undergoing adjuvant chemotherapy in DCISM patients corresponds to a higher risk score according to the nomogram. This might be due to patients at higher risk of relapse being more likely to be selected for chemotherapy. In addition, there are some plausible explanations for why PR status passed the selection process and was kept in the nomogram while ER status did not. Firstly, our study supported the notion that ER positive, PR negative breast cancer is associated with reduced benefits from endocrine therapy[33] and worse clinical outcomes. Secondly, ER positive breast cancers have a higher distant recurrence risk than triple-negative breast cancer[34], so these patients are mostly treated with endocrine therapy which would significantly reduce distant recurrence. Due to the lack of information about endocrine therapy, this therapeutic variable was not included during construction of the nomogram, which may have led to ER status being left out of the nomogram. Prognostic implications of applying different adjuvant treatments are further shown in Figure 4. After propensity score matching, cases treated with chemotherapy had better BCSS as expected, supporting the explanation that patients received chemotherapy had higher risk score was due to clinicopathological factors. When statistically matched, radiotherapy showed no correlation with prognosis, indicating this adjuvant local-regional treatment might contribute more to local control than to BCSS.

There were several limitations in the study. Firstly, the information regarding the HER2/neu status is only available in the SEER database from 2010 onwards. If cases that were diagnosed before 2010 were excluded, the sample size would be dramatically reduced and follow-up time insufficient. Therefore, all cases diagnosed between 1988-2015 were enrolled and HER2 status was not included in the construction of the nomogram. Secondly, the SEER database lacks information about endocrine therapy, so this potential confounding factor could not be analyzed. Thirdly, the retrospective nature of our study may have introduced a certain level of bias in our analysis results. Finally, the sequence of treatment was not considered. Because neither recurrence nor progression is recorded in SEER, we had to treat the therapies as baseline variables instead of time-varying covariates. As a result, it was assumed that the exact treatment combination was determined and given at the time point of diagnosis. Since the exact timing of the treatment is not available, relying on this assumption is necessary to incorporate the therapeutic information into the nomogram.

Conclusion

Our study is the first to develop a nomogram prognostic model specifically for DCISM patients. This model was constructed by using a large population-based training dataset and involved multiple treatment facilities, resulting in a smaller sampling bias. Researchers, clinicians, and patients themselves can easily predict the survival probability for each individual case using the readily available clinical information. This would help providers counsel patients more accurately about their prognosis and determine the best treatment strategy. Additionally, this comprehensive and individualized risk score calculation method may be used as stratification criteria in randomized studies and clinical trials.

Abbreviations

DCISM: Ductal carcinoma in situ with microinvasion; DCIS: Ductal carcinoma in situ; BCSS: breast cancer-specific survival; AUC: area under the receiver operating characteristic curve; ROC: receiver operating characteristic; NCCN: National Comprehensive Cancer Network; AJCC: The American Joint Committee for Cancer; SEER: Surveillance, Epidemiology and End Results; HER2/neu: human epidermal growth factor receptor-2; ICD-O-3: International

Classification of Disease for Oncology, Third Edition; ER: estrogen receptor; PR: progesterone receptor; NSABP: National Surgical Adjuvant Breast and Bowel Project; HR: hazard ratio; CI: confidence interval.

Declarations

Ethics approval and consent to participate: Study data was obtained from the SEER database of the National Cancer Institute, an open access resource for epidemiologic and survival analyses of various cancers. All data is publicly available and de-identified, and therefore exempted from the review of an Institutional Review Board. SEER database data do not require informed consent.

Consent for publication: We have obtained consents to publish this paper from all the participants of this study.

Availability of data and materials: The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing Interests: The authors declare that they have no competing interests.

Funding: This work was supported by the National Natural Science Foundation of China (Grant Numbers: 81902682, 81902303); Shanghai Sailing Program from Science and Technology Commission of Shanghai Municipality (Grant Number: 19YF1408900); Guangdong Basic and Applied Basic Research Foundation (Grant Number: 2020A151501048); Shenzhen Science and Technology Project (Grant Number: JCYJ20190806164216661); and Clinical Research Project of Shenzhen Second People's Hospital (Grant Number: 20203357028).

Authors' contributions: YZZ, YCX, ZHD and GWL designed the research. YZZ and YCX performed the research and analyzed results. YZZ wrote the paper. HBQ, ZZL, GZ, HSJ, SWY and XMW edited the manuscript and provided critical comments. All authors read and approved the final manuscript.

Acknowledgements: The authors would like to acknowledge the helpful comments on this paper received from the reviewers.

References

1. Lakhani SR. WHO Classification of Tumours of the Breast. International Agency for Research on Cancer; 2012.
2. Edge SB, Byrd DR, Carducci MA, Compton CC, Fritz AG, Greene FL, Trotti A. AJCC cancer staging manual, 7th edn. 2010
3. Hoda S, Chiu A, Prasad M, Giri D, Hoda R. Are microinvasion and micrometastasis in breast cancer mountains or molehills? American journal of surgery 2000;180:305-8
4. Adamovich T, Simmons R. Ductal carcinoma in situ with microinvasion. American journal of surgery 2003;186:112-6
5. Bianchi S, Vezzosi V. Microinvasive carcinoma of the breast. Pathology oncology research : POR 2008;14:105-11
6. Shatat L, Gloyeske N, Madan R, O'Neil M, Tawfik O, Fan F. Microinvasive breast carcinoma carries an excellent prognosis regardless of the tumor characteristics. Human pathology 2013;44:2684-9

7. Silver S, Tavassoli F. Mammary ductal carcinoma in situ with microinvasion. Cancer 1998;82:2382-90
8. Padmore R, Fowble B, Hoffman J, Rosser C, Hanlon A, Patchefsky A. Microinvasive breast carcinoma: clinicopathologic analysis of a single institution experience. Cancer 2000;88:1403-9
9. Siegel R, Miller K, Jemal A. Cancer statistics, 2020. CA: a cancer journal for clinicians 2020;70:7-30
10. Pfeiffer R, Webb-Vargas Y, Wheeler W, Gail M. Proportion of U.S. Trends in Breast Cancer Incidence Attributable to Long-term Changes in Risk Factor Distributions. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2018;27:1214-22
11. Gradishar W, Anderson B, Abraham J, Aft R, Agnese D, Allison K, Blair S, Burstein H, Dang C, Elias A, Giordano S, Goetz M, Goldstein L, Isakoff S, Krishnamurthy J, Lyons J, Marcom P, Matro J, Mayer I, Moran M, Mortimer J, O'Regan R, Patel S, Pierce L, Rugo H, Sitapati A, Smith K, Smith M, Soliman H, Stringer-Reasor E, Telli M, Ward J, Young J, Burns J, Kumar R. Breast Cancer, Version 3.2020, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network : JNCCN 2020;18:452-78
12. Gradishar WJ, Anderson BO, Balassanian R, Blair SL, Burstein HJ, Cyr A, Elias AD, Farrar WB, Forero A, Giordano SH, Goetz M, Goldstein LJ, Hudis CA, Isakoff SJ, Marcom PK, Mayer IA, McCormick B, Moran M, Patel SA, Pierce LJ, Reed EC, Salerno KE, Schwartzberg LS, Smith KL, Smith ML, Soliman H, Somlo G, Telli M, Ward JH, Shead DA, Kumar R. Breast Cancer Version 2.2015. Journal of the National Comprehensive Cancer Network : JNCCN 2015;13:448-75
13. Goetz MP, Gradishar WJ, Anderson BO, Abraham J, Aft R, Allison KH, Blair SL, Burstein HJ, Dang C, Elias AD, Farrar WB, Giordano SH, Goldstein LJ, Isakoff SJ, Lyons J, Marcom PK, Mayer IA, Moran MS, Mortimer J, O'Regan RM, Patel SA, Pierce LJ, Reed EC, Rugo HS, Sitapati A, Smith KL, Smith ML, Soliman H, Telli ML, Ward JH, Young JS, Shead DA, Kumar R. NCCN Guidelines Insights: Breast Cancer, Version 3.2018. Journal of the National Comprehensive Cancer Network : JNCCN 2019;17:118-26
14. Parikh R, Haffty B, Lannin D, Moran M. Ductal carcinoma in situ with microinvasion: prognostic implications, long-term outcomes, and role of axillary evaluation. International journal of radiation oncology, biology, physics 2012;82:7-13
15. Li Y, Zhang S, Wei X, Zhang J. The clinical features and management of women with ductal carcinoma in situ with microinvasion: A retrospective Cohort study. International journal of surgery (London, England) 2015;19:91-4
16. Wang W, Zhu W, Du F, Luo Y, Xu B. The Demographic Features, Clinicopathological Characteristics and Cancer-specific Outcomes for Patients with Microinvasive Breast Cancer: A SEER Database Analysis. Scientific reports 2017;7:42045
17. Champion C, Ren Y, Thomas S, Fayanju O, Rosenberger L, Greenup R, Menendez C, Hwang E, Plichta J. DCIS with Microinvasion: Is It In Situ or Invasive Disease? Annals of surgical oncology 2019;26:3124-32
18. Sopik V, Sun P, Narod S. Impact of microinvasion on breast cancer mortality in women with ductal carcinoma in situ. Breast cancer research and treatment 2018;167:787-95
19. Plichta J, Ren Y, Thomas S, Greenup R, Fayanju O, Rosenberger L, Hyslop T, Hwang E. Implications for Breast Cancer Restaging Based on the 8th Edition AJCC Staging Manual. Ann Surg 2020;271:169-76
20. Lee S, Sohn G, Kim J, Chung I, Lee J, Kim H, Ko B, Son B, Ahn S. A retrospective prognostic evaluation analysis using the 8th edition of the American Joint Committee on Cancer staging system for breast cancer. Breast cancer research and treatment 2018;169:257-66

21. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. The Lancet Oncology 2015;16:e173-e80
22. Iasonos A, Schrag D, Raj G, Panageas K. How to build and interpret a nomogram for cancer prognosis. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2008;26:1364-70
23. Xiong Z, Deng G, Huang X, Li X, Xie X, Wang J, Shuang Z, Wang X. Score for the Survival Probability in Metastasis Breast Cancer: A Nomogram-Based Risk Assessment Model. Cancer research and treatment : official journal of Korean Cancer Association 2018;50:1260-9
24. Lei Z, Li J, Wu D, Xia Y, Wang Q, Si A, Wang K, Wan X, Lau W, Wu M, Shen F. Nomogram for Preoperative Estimation of Microvascular Invasion Risk in Hepatitis B Virus-Related Hepatocellular Carcinoma Within the Milan Criteria. JAMA surgery 2016;151:356-63
25. Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36
26. Kapoor N, Shamonki J, Sim M, Chung C, Giuliano A. Impact of multifocality and lymph node metastasis on the prognosis and management of microinvasive breast cancer. Annals of surgical oncology 2013;20:2576-81
27. Mamtani A, Nakhlis F, Downs-Canner S, Zabor EC, Morrow M, King TA, Van Zee KJ. Impact of Age on Locoregional and Distant Recurrence After Mastectomy for Ductal Carcinoma In Situ With or Without Microinvasion. Annals of surgical oncology 2019;26:4264-71
28. Si J, Guo R, Pan H, Lu X, Guo Z, Han C, Xue L, Xing D, Wu W, Chen C. In Situ Multiple Microinvasion Foci in Ductal Carcinoma Is Associated With an Increased Risk of Recurrence and Worse Survival Outcome. Frontiers in oncology 2020;10:607502
29. Foulkes WD, Reis-Filho JS, Narod SA. Tumor size and survival in breast cancer--a reappraisal. Nature reviews Clinical oncology 2010;7:348-53
30. Bernards R, Weinberg RA. A progression puzzle. Nature 2002;418:823
31. Zheng Y-Z, Wang X-M, Fan L, Shao Z-M. Breast Cancer-Specific Mortality in Small-Sized Tumor with Stage IV Breast Cancer: A Population-Based Study. The oncologist 2020;25:1–10
32. Wrubel E, Natwick R, Wright GP. Breast-Conserving Therapy is Associated with Improved Survival Compared with Mastectomy for Early-Stage Breast Cancer: A Propensity Score Matched Comparison Using the National Cancer Database. Annals of surgical oncology 2021 Feb;28:914-9
33. Liu X-Y, Ma D, Xu X-E, Jin X, Yu K-D, Jiang Y-Z, Shao Z-M. Genomic Landscape and Endocrine-Resistant Subgroup in Estrogen Receptor-Positive, Progesterone Receptor-Negative, and HER2-Negative Breast Cancer. Theranostics 2018;8:6386-99
34. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. The New England journal of medicine 2010;363:1938-48

Tables

Table 1. The individual characteristics of variables involved in the study.

Characteristics	Training set		Validating set		*P*[a] value
	No. of patients	%	No. of patients	%	
Year of diagnosis					0.157
1988-2001	391	10.3	183	11.2	
2002-2015	3415	89.7	1449	88.8	
Age					0.213
≤40	201	5.3	71	4.4	
40-50	915	24.0	402	24.6	
51-70	2048	53.8	863	52.9	
≥70	642	16.9	296	18.1	
Marital status					0.199
Married	1284	33.7	542	33.2	
Not married[b]	2357	61.9	1036	63.5	
Unknown	165	4.3	54	3.3	
Race					0.199
Caucasian	2894	76.0	1231	75.4	
African American	445	11.7	188	11.5	
American Indian/Alaskan native, or Asian/Pacific Islander	467	12.3	213	13.1	
AJCC N stage					0.213
N0	3511	92.2	1527	93.6	
N1	248	6.5	94	5.8	
N2	32	0.8	7	0.4	
N3	15	0.4	4	0.2	
Primary site					0.242
Upper-inner quadrant of breast	375	9.9	137	8.4	
Lower-inner quadrant of breast	249	6.5	122	7.5	
Upper-outer quadrant of breast or axillary tail of breast	1341	35.2	616	37.7	
Lower-outer quadrant of breast	268	7.0	106	6.5	
Nipple or central portion of breast	217	5.7	106	6.5	

Overlapping lesion of breast	1356	35.6	545	33.4	
Laterality					0.157
Right-origin of primary	1900	49.9	800	49.0	
Left-origin of primary	1906	50.1	832	51.0	
Grade					0.199
I	811	21.3	344	21.1	
II	1478	38.8	693	42.5	
III and undifferentiated	1517	39.9	595	36.5	
ER status					0.157
Positive	2685	70.5	1181	72.4	
Negative	1121	29.5	451	27.6	
PR					0.157
Positive	2177	57.2	971	59.5	
Negative	1629	42.8	661	40.5	
Surgery					0.199
No surgery	23	0.6	16	1.0	
Lumpectomy	2155	56.6	940	57.6	
Mastectomy	1628	42.8	676	41.4	
Chemotherapy					0.157
No	3392	89.1	1499	91.9	
Yes	412	10.8	133	8.1	
Radiation					0.157
No	1998	52.5	827	50.7	
Yes	1808	47.5	805	49.3	

Abbreviations: AJCC, American Joint Committee on Cancer; ER, estrogen receptor; PR, progesterone receptor.

[a]*P* values were calculated by chi-square test.

[b]Includes divorced, separated, single (never married), and widowed.

Table 2. Univariate and multivariate analyses of breast cancer-specific mortality.

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Year of diagnosis				
1988-2001	References		References	
2002-2015	0.445(0.300,0.65)	**<0.0001**	0.559(0.374,0.836)	**0.005**
Age				
<40	References		References	
40-50	0.352(0.189,0.656)	**0.001**	0.411(0.217,0.780)	**0.006**
51-70	0.345(0.197,0.606)	**<0.0001**	0.534(0.296,0.964)	**0.037**
>70	0.753(0.414,1.371)	0.354	1.401(0.716,2.741)	0.324
Marital status				
Not married[a]	References		References	
Married	0.671(0.469,0.961)	**0.029**	0.786(0.533,1.158)	0.223
Unknown	0.745(0.265,2.029)	0.551	0.568(0.201,1.603)	0.285
Race				
Caucasian	References		References	
African American	2.148(1.396,3.305)	**0.001**	1.686(1.061,2.679)	**0.027**
American Indian/Alaskan native, or Asian/Pacific Islander	0.722(0.376,1.389)	0.330	0.753(0.388,1.460)	0.401
AJCC N stage				
N0	References		References	
N1	5.144(3.419,7.740)	**<0.0001**	2.690(1.565,4.624)	**<0.0001**
N2	7.706(3.363,17.658)	**<0.0001**	4.290(1.644,11.193)	**0.003**
N3	18.857(7.641,46.534)	**<0.0001**	9.321(3.226,26.936)	**<0.0001**
Primary site				
Upper-outer quadrant of breast or axillary tail of breast	References		References	
Upper-inner quadrant of breast	1.301(0.678,2.495)	0.428	1.234(0.628,2.427)	0.542
Lower-inner quadrant of breast	0.563(0.201,1.579)	0.275	0.622(0.22,1.755)	0.369
Lower-outer quadrant of breast	0.840(0.355,1.991)	0.693	0.864(0.363,2.059)	0.742
Nipple or central portion of breast	1.258(0.586,2.702)	0.556	1.1(0.508,2.385)	0.808

Overlapping lesion of breast	1.624(1.077,2.449)	0.210	1.396(0.916,2.126)	0.121
Laterality				
Left-origin of primary	References		References	
Right-origin of primary	0.857(0.604,1.216)	0.387	0.745(0.520,1.066)	0.108
Grade				
I	References		References	
II	1.296(0.739,2.273)	0.366	1.136(0.638,2.025)	0.664
III and undifferentiated	1.773(1.040,3.024)	**0.036**	1.229(0.685,2.203)	0.49
ER status				
Negative	References		References	
Positive	0.739(0.516,1.059)	0.100	1.491(0.92,2.417)	0.105
PR				
Negative	References		References	
Positive	0.583(0.410,0.830)	**0.003**	0.503(0.316,0.802)	**0.004**
Surgery				
No surgery	References		References	
Lumpectomy	0.057(0.026,0.125)	**<0.0001**	0.11(0.042,0.284)	**<0.0001**
Mastectomy	0.106(0.049,0.232)	**<0.0001**	0.12(0.051,0.283)	**<0.0001**
Chemotherapy				
No	References		References	
Yes	4.466(3.100,6.435)	**<0.0001**	2.205(1.293,3.761)	**0.004**
Radiation				
No	References		References	
Yes	0.658(0.459,0.942)	**0.022**	0.83(0.508,1.358)	0.459

Abbreviations: HR, hazard ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor.
[a]Includes divorced, separated, single (never married), and widowed. Notes: Significant results were bolded.

Figures

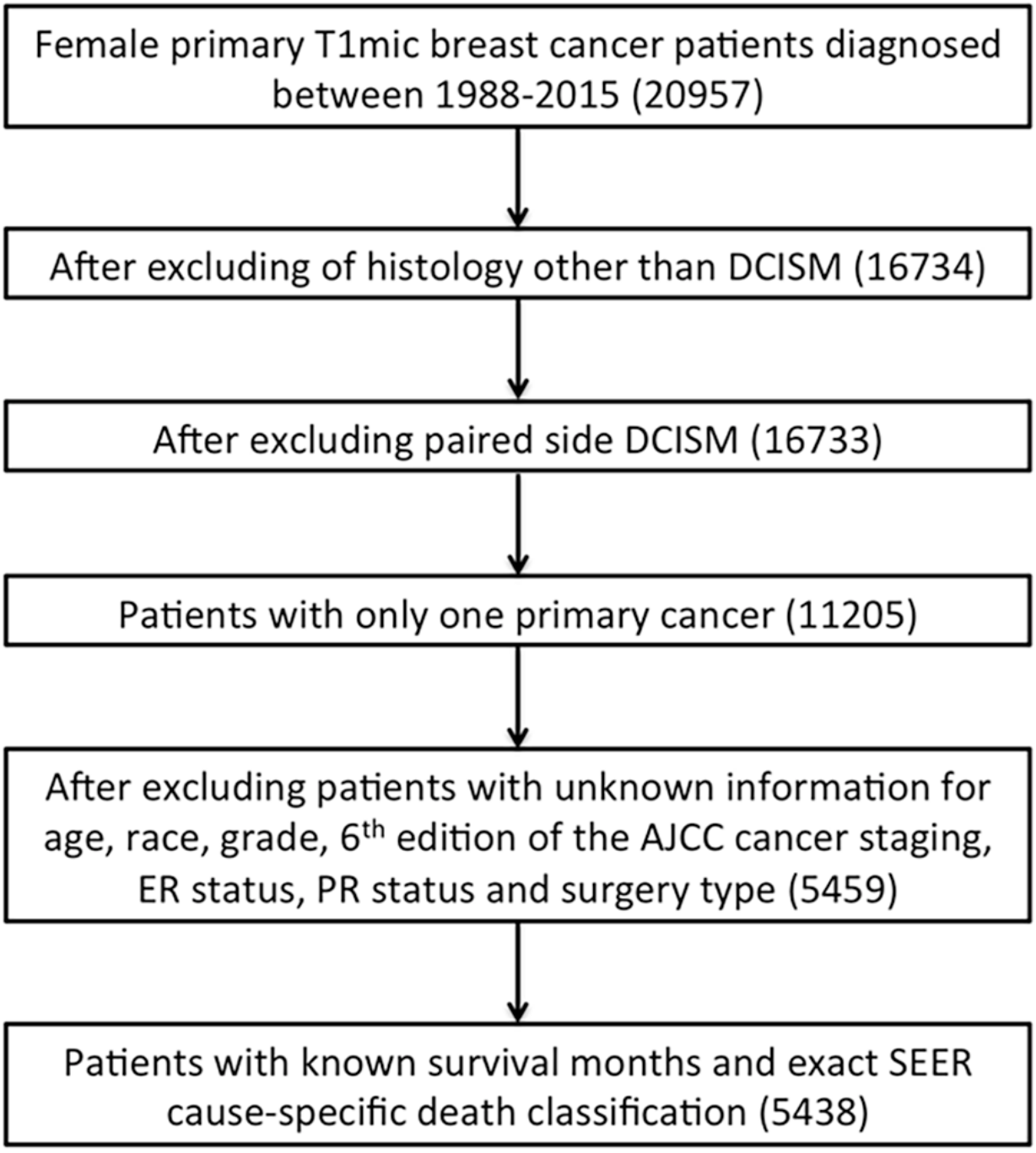

Figure 1

Flowchart of the case selection process in the study.

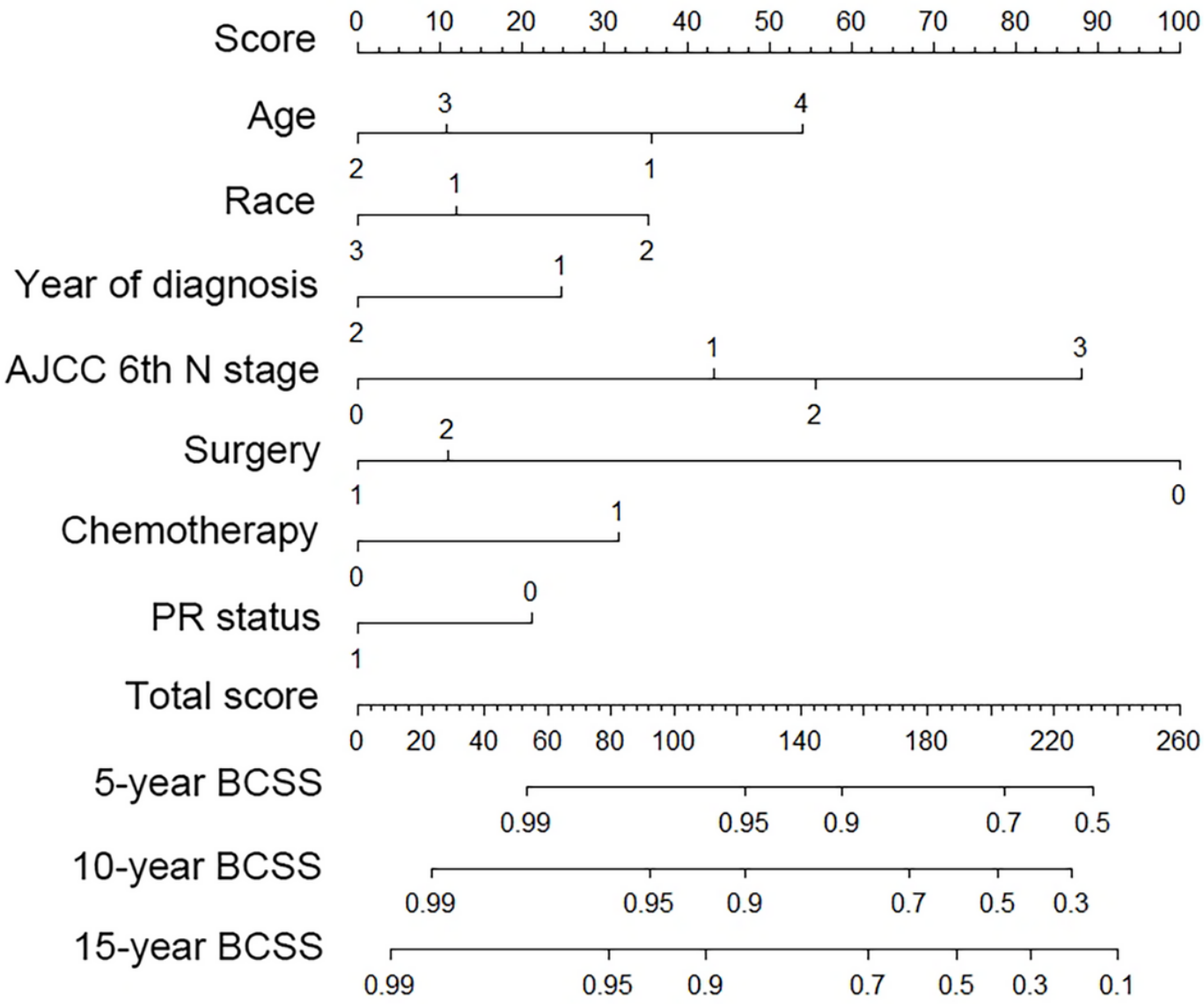

Figure 2

Nomogram to calculate risk score and predict 5-year, 10-year, and 15-year BCSS probability. By summing the points identified on the top scale for each independent variable and drawing a vertical line from the total points scale to the 5-year, 10-year, and 15-year BCSS, the corresponding survival probability can be obtained. Age, 1 = under 40 years, 2 = 41-50 years, 3 = 51-70 years, 4 = over 70 years; Race, 1 = Caucasian, 2 = African American, 3 = American Indian/Alaska Native or Asian/Pacific Islander; Year of diagnosis, 1 = 1988-2001, 2 = 2002-2015; AJCC 6th N stage, 0 = N0 stage, 1 = N1 stage, 2 = N2 stage, 3 = N3 stage; Surgery, 0 = no surgery, 1 = lumpectomy, 2 = mastectomy; Chemotherapy, 0 = no, 1 = yes; PR status, 0 = negative, 1 = positive. BCSS: breast cancer-specific survival.

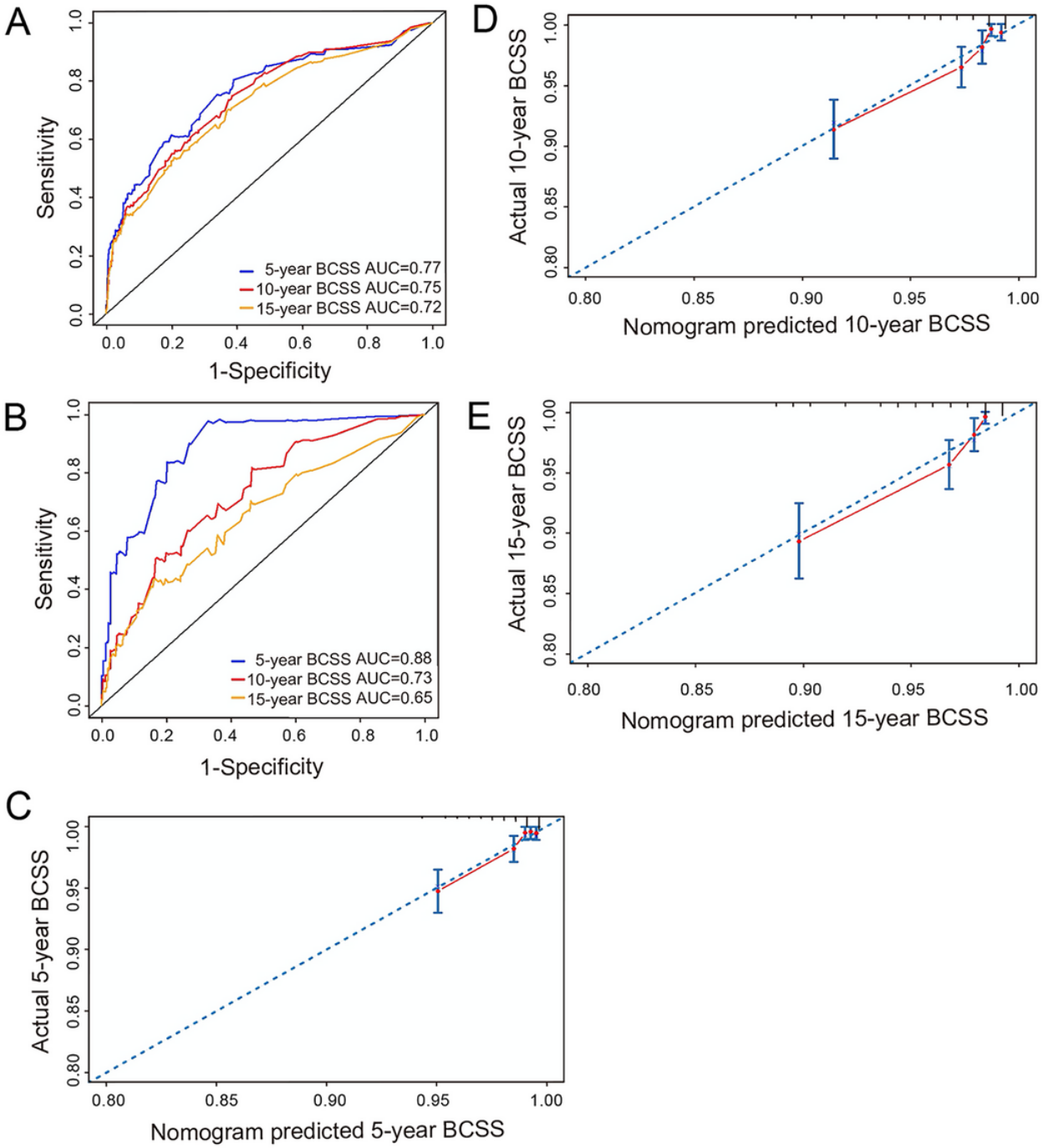

Figure 3

ROC curves and calibration plots for predicting BCSS. ROC curves of the nomogram predicting prognosis in the training set (A) and the validating set (B). Calibration curves comparing predicted and actual BCSS proportions at 5-year (C), 10-year (D), and 15-years (E), separately. Each point in the plot refers to a group of patients, with the nomogram predicted probability of survival shown on x axis and actual survival proportion shown on y axis. Distributions of predicted survival probabilities are plotted at the top. Error bars represent 95% confidence intervals. ROC: receiver operating characteristic curve; AUC: area under the ROC curve; BCSS: breast cancer-specific survival.

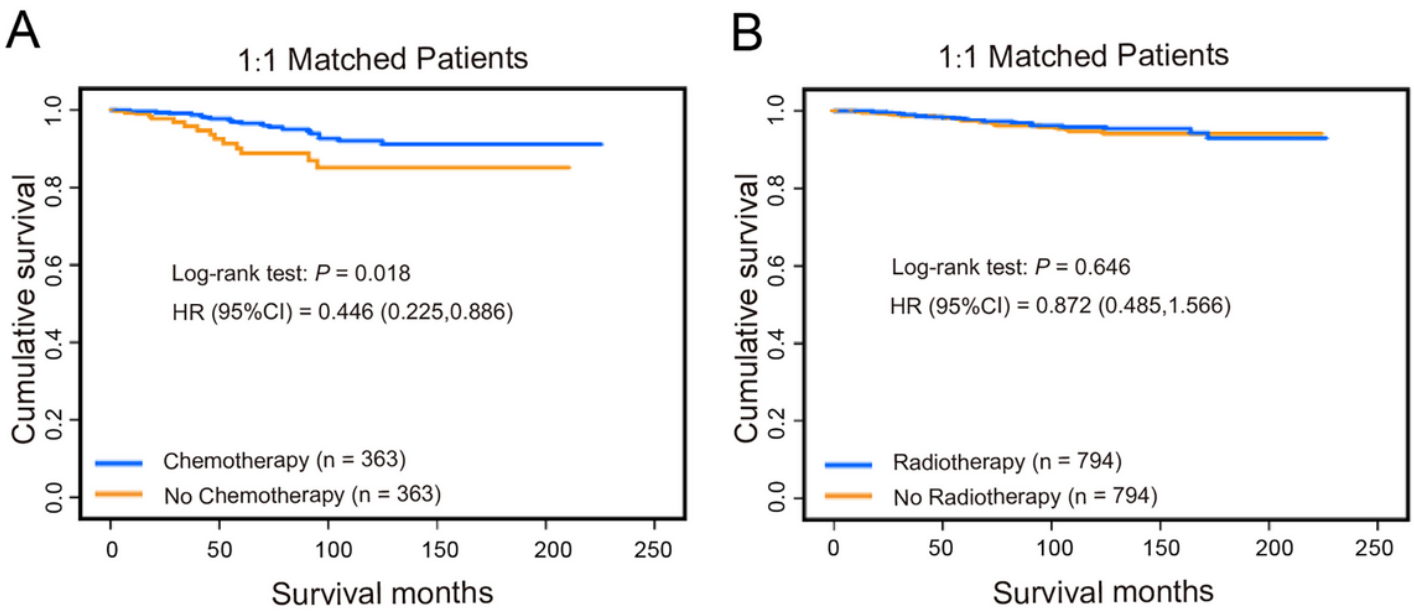

Figure 4

The survival curves for DCISM patients with and without chemo¬therapy and radiotherapy after 1:1 matching. A. Kaplan-Meier curve depicting the association between chemotherapy and breast DCISM (log-rank test, P = 0.018). B. Kaplan-Meier curve depicting the association between radiotherapy and breast DCISM (log-rank test, P = 0.872). DCISM: Ductal carcinoma in situ with microinvasion (DCISM).