A Nomogram Individually Predicts the Overall Survival of Patients with Breast Cancer after Surgery：a Retrospective Study in the SEER Database and China

Yufen Zheng
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University
Minya Jin
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University
Shiyong Chen
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University
Yijun Feng
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University
Juping Du
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University
Jun Li
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University
Jing Wang (wjing@enzemed.com)
Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University

Article

Keywords: Breast cancer, nomogram, overall survival, prognosis, SEER Database

Posted Date: May 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1306867/v2

Abstract

Background

Patients with breast cancer have a poor prognosis. We want to construct a more elaborative and validate nomograms for predicting overall survival in patients with breast cancer.

Methods

A total of 68363 breast cancer patients who underwent surgery between 2011 and 2015 were recruited from the Surveillance Epidemiology, and End Results (SEER) database. After eliminating lacking clinical information, 60445 eligible breast cancer patients were randomly divided into the training corhort(n = 42327) and the internal validation corhort(n = 18118) in a ratio of 7:3. The endpoint of this study was overall survival(OS). Multivariate Cox proportional hazards regression models was performed to identify independent risk factors of OS in the training corhort, and then the nomogram was constructed. The nomogram predictive performance was evaluated by the Harrell's concordance index (C-index), the time-dependent receiver operating characteristics (ROC) curve (AUC), calibration curve, decision curve analysis(DCA) and clinical impact curve. Moreover, the nomogram was verified by the internal validation corhort and external validation corhort.

Results

Age, gender, grade, 7th AJCC stage, ER status, PR status, Her-2 Status, breast subtype were found to be independent risk factor of OS($P < 0.05$). The nomogram integrating these eight factors was constructed and proved excellent discrimination capability in the training corhort(C-index, 0.724 (95%CI, 0.716–0.732)), which was demonstrated in the internal validation corhort(C-index, 0.717 (95% CI, 0.705–0.729) and external validation corhort(C-index, 0.793 (95% CI, 0.724–0.862)). Calibration curve for the probability of 1-, 3- and 5-year OS demonstated good concordance between nomogram prediction and actual observed results in both the training and validation corhort. Besides, the DCA and clinical impact curve indicated the clinical usefulness of our constructed nomogram.

Conclusions

We developed a nomogram that integrate clinicopathological variables, which can precisely predict the 1-, 3- and 5-year OS of breast cancer patients after surgery. Validation uncovered preeminent discrimination power for the nomogram, indicating that it presents satisfactory clinical application. Therefore, the nomogram can help clinicians in formulating the suitable therapy strategies for individual patients.

Background

Breast cancer is one of the most common malignancies in the worldwide and the second leading cause of cancer death in female patients[1, 2]. The incidence of breast cancer has been increasing gramatically in recent decades, and the exact pathogenesis remains not clearly known yet. With the continuous progress of various treatment modalities in latest years, but the prognosis of breast cancer is still remains poor.

The American Joint Committee on Cancer (AJCC) TNM staging system is globally recognized and widely used to predict disease progression and design effective therapeutic strategies of breast cancer patients[3, 4], but it is largely influenced by a variety of conclusive clinicopathological characteristics, for example age, gender, histological differentiation and breast subtype, which may also have a favorable impact on survival in patients[5, 6]. If the prognosis could be accurately predicted in breast cancer patients after resection, comprehensive scientific treatment would be taken timely to high-risk patients to increase the survival and decrease the mortality.

The nomogram is considered as a useful and reliable clinical tool to help clinicians and patients estimate the overall survival probability and make personalized decisions by incorporating vital clinical indicators to predict the survival outcome of certain cancers individually[7, 8].The good accuracy and simplicity of the prediction of the nomogram make it a new standard to guide the treatment of cancer patients[9].

In this study, our aim is to build a more elaborative nomogram to predict the overall survival rates of breast cancer patients with relatively large cohort in 1-, 3- and 5-year, based the surveillance, epidemiology, and end results (SEER) database.

Methods

Study population selection and design

We acquired the study population from the SEER (Version 8.3.8)program records of the National Cancer Institute (https://seer.cancer.gov/). The SEER database can provide more information about demographic, tumor location, tumor size and morphology, diagnosis stage, treatment, prognosis and so on of millions of cancer patients, which provides excellent data for clinical research of tumors. The 1975-2016 year SEER (submitted in November 2018)) became useable on April 2019.

The inclusion criteria were as follows: (1) patients aged at diagnosis≥14 years who underwent surgery at the site of the primary tumor, (2) pathology histological diagnosis was breast cancer, (3) year of diagnosis between 2011-2015, (4) complete clinicopathological follow-up data. The exclusion criteria were(1) patients diagnosed according to exfoliative cytology, (2) patients with unkown 7th AJCC stage, ER status, PR status, Her2 status, breast subtype, survival time and follow-up vital status. Ultimately, the data of 60445 patients were enrolled and analysed in this study from the SEER database. Furthermore, the eligible breast cancer patients were randomly splited into two groups in a ratio of 7:3, the training corhort(n=42327) and internal validation corhort(n=18118). In addition, a chinese corhort (n=332) from

Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University was used for external validation. Our study procedure is shown in *Figure 1*.

Extraction of demographic and clinicopathological characteristics from SEER database, containing age at diagnosis, gender, year of diagnosis, tumor site, behavior recode, grade, histology diagnostic confirmation, 7th AJCC stage, ER(estrogen receptor) status, PR(progesterone receptor) status, Her2(human epidermal growth factor receptor-2) status, breast subtype, survival months and follow-up vital status. The chinese corhort follow-ups ended in April 29, 2020 or the date of patient death, whichever came first. The overall survival time(OS) was defined as the time from diagnosis to death for any reason or last follow-up.

The analysis of data from SEER is not limited by medical ethics review and does not require informed consent. The external validation cohort was approved by the hospital ethics committee, but because it was a retrospective study, informed consent was not required(project number: K20210809) . All procedures conducted in studies involving human participants met the 1964 declaration of Helsinki and its subsequent amendments or similar ethical standards.

Statistical Analyses

Categorical variables were described as frequency and proportion. Continuous variables were described as mean (standard deviation, SD) or median (interquartile ranges,IQR). In order to compare the basic characteristics between training corhort and validation corhort, chi-square test or Fisher's exact test was used for categorical variables, student's t-test was used for normal distribution variables, and Mann-Whitney U-test was used for non normal distribution variables. Univariate and multivariate Cox proportional risk regression model was used to recognize independent risk factors of OS in the training corhort($P<0.05$). The survival differences were compared between groups using Kalan-Meier curves. These independent risk factors were used to construct nomogram for breast cancer patients. The nomogram was validated internally(training corhort) and externally(internal validation cohort and external validation cohort). The calibration curves were used to compare the correlation between actual and predicted survival outcomes[9]. Both the Harrell’s concordance index (C-index) and area under the time-dependent receiver operating characteristics (ROC) curve (AUC) can be used to assess the discrimination accuracy of the nomogram[10]. The decision curve analyses(DCA) and clinical impact curve were used to evaluate the clinical usefulness and benefits of the predictive model[11] The SPSS 22.0 and the R software (version 4.0.3) were used for All statistical analysis.Two tailed test P-value less than 0.05 was considered statistically significant.

Results

Study population characteristics

The demographic and clinicopathological characteristics of the breast cancer patients are listed in *Table 1*. The median follow-up time of all patients were 48.0 months, and 6091(10.1%) breast cancer patients

die of all of cause. On the basis of age at diagnosis, all patients were be segmented into two groups: <60 year(42.9%), ≥60 year(57.1%). Most of patients presented with pathological moderately differentiated grade II (44.2%), 7th AJCC stage I(54.4%) at the time of diagnosis. Of these, 45677(75.5%) patients of breast subtype were Luminal A.

Independent prognostic factors of OS

In the training corhort, the results of univariate and multivariate analysis are shown in *Table 2*. The multivariate analysis revealed that age, gender, grade, 7 th AJCC stage, ER status, PR status, Her-2 status, breast subtype were found to be independent risk factor of OS(P<0.05). It's more intuitionistic to present the results as a forest plot (*Figure 2*).

Figure 3 displays the Kaplan-Meier survival curve in training corhort. The patients with above 60 years old, ER negative, PR negative, Her-2 negative and triple negative had shorter survival time(P<0.05). The grade IV and 7 th AJCC stage IV group had the worst prognosis(P<0.0001).

Nomogram Construction

Based on the multivariate Cox regressions analyses results, a nomogram was constructed (*Figure 4*) for predicting OS in training corhort. The nomograms indicated that 7th AJCC stage contributed most to OS. Just to make it easier to use the nomogram, every variable was allocated a score that a vertical line is drawn upward to ascertain the number of points of each variable value. Add these numbers and find the corresponding position on the total point axis, a straight line is drawn downward to the survival axis to ascertain the possibility of OS. For example, the patient was less than 60 years old, grade , 7th AJCC stage , ER status negative, PR status negative, Her-2 status positive and the breast subtype was luninal B. Her total points were 13.2, and the probability of 1-year OS survival, 3-year OS survival, 5-year OS survival was more than 95%, 87%, 75%, respectively, on the basis of the nomogram.

Nomogram Validation

The discrimination performance of nomogram was 0.724 (95%CI, 0.716-0.732), which was confirmed in the internal validation corhort (C-index, 0.717 (95%CI, 0.705-0.729) and external validation corhort(C-index, 0.793 (95%CI, 0.724-0.862), respectively. Moreover, the nomogram proved better differentiated ability for OS compared with the 7th AJCC stage in both the training corhort(AUC:0.726, 0.634, respectively, *Figure 5A*), internal validation corhort(AUC:0.716, 0.613, respectively, *Figure 5B*) and external validation cohort(AUC:0.716, 0.613, respectively, *Figure 5C*).

The calibration curve for predicting 1-, 3-, or 5-years OS present a good concordance of the nomogram predicted probability with actual observations in both the training corhort, internal and external validation corhort(*Figure 6*). Ultimately, the DCA curves was carried out to indicated the clinical net benefits of the nomogram with that of the 7th AJCC stage, grade(*Figure 7A,7B,7C*). The nomogram predict the OS probabilities showed largest net benefits in the training, internal and external validation corhor(*Figure 7D,7E,7F*).

Discussion

In this study, we construct a nomogram to predict the OS of breast cancer patients after surgery, which integrated age at diagnosis, gender, grade, 7 th AJCC stage, ER status, PR status, Her-2 status, and breast subtype. Furthermore, the nomogram was validated using the internal and external validation cohort. This nomogram was more significantly predictive(AUC: 0.735) than the 7th AJCC stage(AUC: 0.634). In addition, the calibration curve showed that the predicted 1-, 3- and 5-year OS closely related to the actual observations, whether it is in training, internal and external validation corhort. Similarly, the DCA curve manifested it had largest net benefits and promising clinical applicability.

According to the results of multivariate cox regression analysis, eight clinicopathological characteristics were found to be independent dangerous factors, containing age at diagnosis, gender, grade, 7th AJCC stage, ER status, PR status, Her-2 status and breast subtype. The nomogram reveals that the 7th AJCC stage is the most important variable affecting OS, which is mainly due to the 7th AJCC stage, including tumor size, lymph node metastasis and distant metastasis, which are very important factors affecting the prognosis of breast cancer[12, 13]. Our results indicated that patients who were older than 60 years, the older they were, the worse the prognosis was. As similar with our consequences, age was proved to be a significant association with OS in several other studies[14–16]. One possible reason for this that OS might be influenced by age not only associated with the clinical course of a disease, but also with age-related complications[17]. Moreover, aging might facilitate the growth of tumor cells by suppressing the immune system[18]. As can be seen from the nomogram, breast cancer patients who are negative for ER and PR had a poor prognosis, which is in accordance with other findings[19, 20]. Histological grade showed undifferentiated tumors and triple negative breast cancer, which has been recognized as an index of a poor prognosis in breast cancer patients[16, 21].

The nomogram is a imaging exhibition of a statistical prediction model that provides survival probability of a particular outcome[22, 23]. Therefore, the parameters should be readily retrievable and gaugeable. What's more, the nomogram has enough discriminant capability, and the prediction is in good agreement with the actual observation. In terms of prognosis, nomogram of other tumors has also been proved to be economical and practical. Fang et al.[24]established the nomogram by combining age, tumour size, differentiation, N stage, M stage with tumor location to precisely predict OS of gastroenteropancreatic neuroendocrine neoplasms. Kong et al.[6] integrating age at diagnosis,T stage, N stage, and M stage to construct a credible and powerful nomogram to predict prognosis for adrenocortical carcinoma patients after surgery. There is growing evidence has shown that the nomogram presents a better predictive capacity than the traditional AJCC TNM stage in numerous tumors[24–26]. In comparison with the extensively used AJCC TNM stage, our nomogram is not merely simple and convenient, but also provides a accurate individual prognosis for different patients. Therefore, the nomogram will help clinicians to estimate personalized survival probability and make with optimize therapeutic schedule.

In addition to the reliable data sources of nomogram, our study has several advantages. Firstly, the clinicopathological characteristic of breast cancer patients that we collected from SEER database was abundant and comprehensive, therefore assuring construct the accurate and reliable prognostic nomogram. Secondly, the nomogram presents superior discrimination ability in predicting OS compared with 7th AJCC stage. The effectiveness and practicability of the nomogram was verified by the internal and external validation corhort. Finally, this research employed eight clinicopathological variables that are easily to obtain and extensively used in clinical practice, which bring simple and convenient for using of the nomogram.

This study had the following shortcomings. First, this is a retrospective study from the SEER database, selection bias was hardly inevitable. Second, we can not get more clinical information from SEER database, such as vascular invasion, radiotherapy, chemotherapy and laboratory data, which, if contained, the sensitivity and specificity of nomogram can be improved. Third, the lack of clinical data and follow-up may affect the discrimination and prediction ability of nomogram. We should design a prospective experiment to further verify the nomogram in future.

Conclusions

In summary, construction of nomogram to predict the postoperative OS in breast cancer patients. Our user-friendly nomogram, an easy-to-use tool for risk evaluation and survival prediction in personalized breast cancer patients, which can precisely and efficiently provide individualized counseling, timely surveillance, and clinical evaluation.

Abbreviations

SEER: Surveillance Epidemiology, and End Results, OS: overall survival, C-index: Harrell's concordance index, AUC: time-dependent receiver operating characteristics curve, DCA: decision curve analysis, AJCC: American Joint Committee on Cancer, ER: estrogen receptor, PR: progesterone receptor, Her-2:human epidermal growth factor receptor-2.

Declarations

Acknowledgments

We would like to thank all the staff members of the National Cancer Institute who work with the SEER program.

Author contributions

Yufen Zheng: Designed the project, collected datas and performed data analysis, worte the manuscript. Minya Jin and Shiyong Chen: Designed the project, performed data analysis and worte the manuscript. Yijun Feng and Juping Du: Performed data analysis and worte the manuscript. Jun Li: Supervised the

project, administrative support and worte the manuscript. Jing Wang: Designed the project, supervised the project and worte the manuscript. All authors read and approved the final manusript.

Funding

This project is supported by the Medical Science and Technology Project of Zhejiang Province (2021KY394).

Availability of data and materials

The data used in the study is publicly available and is obtained from the corresponding author.

Ethics approval and consent to participate

The analysis of data from SEER is not limited by medical ethics review and does not require informed consent. The external validation cohort was approved by the Taizhou Hospital of Zhejiang Province ethics committee, but because it was a retrospective study, informed consent was not required. All procedures conducted in studies involving human participants met the 1964 declaration of Helsinki and its subsequent amendments or similar ethical standards.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University, Linhai 317000, Zhejiang Province, China.

References

1. Cardoso F, Harbeck N, Fallowfield L, et al. Locally recurrent or metastatic breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology. 2012; 23: vii11-vii9.
2. Lee SB, Sohn G, Kim J, et al. 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.
3. Plichta JK, Ren Y, Thomas SM, et al. Implications for breast cancer restaging based on the 8th edition AJCC staging manual. Annals of surgery. 2020; 271: 169–76.
4. Fang C, Wang W, Deng J-Y, et al. Proposal and validation of a modified staging system to improve the prognosis predictive performance of the 8th AJCC/UICC pTNM staging system for gastric adenocarcinoma: a multicenter study with external validation. Cancer communications. 2018; 38: 67.

5. Yao J-J, Zhou G-Q, Wang Y-Q, et al. Prognostic values of the integrated model incorporating the volume of metastatic regional cervical lymph node and pretreatment serum Epstein–Barr virus DNA copy number in predicting distant metastasis in patients with N1 nasopharyngeal carcinoma. Chinese journal of cancer. 2017; 36: 1–7.
6. Kong J, Zheng J, Cai J, et al. A nomogram for individualized estimation of survival among adult patients with adrenocortical carcinoma after surgery: a retrospective analysis and multicenter validation study. Cancer communications. 2019; 39: 80.
7. Fu YP, Yi Y, Huang JL, et al. Prognostic nomograms stratify survival of patients with hepatocellular carcinoma without portal vein tumor thrombosis after curative resection. The oncologist. 2017; 22: 561.
8. Dong F, Shen Y, Gao F, et al. Nomograms to predict individual prognosis of patients with primary small cell carcinoma of the bladder. Journal of Cancer. 2018; 9: 1152.
9. Coutant C, Olivier C, Lambaudie E, et al. Comparison of models to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a prospective multicenter study. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2009; 27: 2800–8.
10. Wang W, Sun Z, Deng J-Y, et al. A novel nomogram individually predicting disease-specific survival after D2 gastrectomy for advanced gastric cancer. Cancer communications. 2018; 38: 1–9.
11. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Medical Decision Making. 2006; 26: 565–74.
12. Lin S, Mo H, Li Y, et al. Development and validation of a nomogram for predicting survival of advanced breast cancer patients in China. The Breast. 2020; 53: 172–80.
13. Diao Jd, Ma Lx, Sun My, et al. Construction and validation of a nomogram to predict overall survival in patients with inflammatory breast cancer. Cancer Medicine. 2019; 8: 5600–8.
14. X. Cui, H. Zhu, and J. Huang, Nomogram for Predicting Lymph Node Involvement in Triple-Negative Breast Cancer. Frontiers in oncology 10 (2020) 2681.
15. Cui X, Zhu H, Huang J. Nomogram for Predicting Lymph Node Involvement in Triple-Negative Breast Cancer. Frontiers in oncology. 2020; 10: 2681.
16. Liu D, Wu J, Lin C, et al. Breast Subtypes and Prognosis of Breast Cancer Patients With Initial Bone Metastasis: A Population-Based Study. Frontiers in oncology. 2020; 10: 2699.
17. Fang C, Wang W, Zhang Y, et al. Clinicopathologic characteristics and prognosis of gastroenteropancreatic neuroendocrine neoplasms: a multicenter study in South China. Cancer communications. 2017; 36: 1–9.
18. Yancik R, Wesley MN, Ries LA, et al. Effect of age and comorbidity in postmenopausal breast cancer patients aged 55 years and older. Jama. 2001; 285: 885–92.
19. Gong Y, Liu Y-R, Ji P, et al. Impact of molecular subtypes on metastatic breast cancer patients: a SEER population-based study. Scientific reports. 2017; 7: 45411.

20. Leone JP, Leone J, Zwenger AO, et al. Prognostic factors and survival according to tumour subtype in women presenting with breast cancer brain metastases at initial diagnosis. European Journal of Cancer. 2017; 74: 17–25.
21. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the united States of America. 2003; 100: 8418-23.
22. Kattan MW, Leung DH, Brennan MF. Postoperative nomogram for 12-year sarcoma-specific death. Journal of clinical oncology. 2002; 20: 791–6.
23. See WA. Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer: International Bladder Cancer Nomogram Consortium, Bochner BH, Kattan MW, Vora KC, Department of Urology, Memorial Sloan-Kettering Cancer Center, Kimmel Center for Prostate and Urologic Tumors, New York, NY. Urologic Oncology: Seminars and Original Investigations: Elsevier; 2007. p. 275.
24. Fang C, Wang W, Feng X, et al. Nomogram individually predicts the overall survival of patients with gastroenteropancreatic neuroendocrine neoplasms. British journal of cancer. 2017; 117: 1544–50.
25. Chen S, Liu Y, Yang J, et al. Development and validation of a nomogram for predicting survival in male patients with breast cancer. Frontiers in oncology. 2019; 9: 361.
26. Li C, Yang J, Zheng S, et al. Establishment and Validation of a Nomogram for Tonsil Squamous Cell Carcinoma: A Retrospective Study Based on the SEER Database. Cancer Control. 2020; 27: 1073274820960481.

Tables

Table 1. The demographic and clinicopathological characteristics in the SEER database

Variable	Total cohort n=60445	Training cohort n=42327	Internal validation cohort n=18118	External validation cohort n=332
Age (Years)				
60	25965(42.9%)	18236(43.0%)	7729(42.6%)	268(80.7%)
≥60	34480(57.1%)	24091(57.0%)	10389(57.4%)	64(19.3%)
Grade				
Grade I	15976(26.4%)	11068(26.1%)	4908(27.1%)	13(3.9%)
Grade II	26766(44.2%)	18886(44.6%)	7880(43.5%)	164(49.4%)
Grade III	17613(29.1%)	12315(29.0%)	5298(29.2%)	155(46.7%)
Grade IV	90(0.1%)	58(0.1%)	32(0.2%)	0(0.0%)
7th AJCC T stage				
T0	41(0.1%)	29(0.1%)	12(0.1%)	0(0.0%)
T1	38854(64.3%)	27277(64.4%)	11577 (63.9%)	133(40.1%)
T2	17156(28.3%)	11943(28.2%)	5213(28.8%)	183(55.1%)
T3	3212(5.3%)	2253(5.3%	959(5.3%)	16(4.8%)
T4	1182(2.0%)	825(2.0%)	357(2.0%)	0(0.0%)
7th AJCC N stage				
N0	44061(72.9%)	30938(73.1%)	13123(72.4%)	157(47.3%)
N1	12399(20.5%)	15187(35.2%)	3779(20.9%)	106(31.9%)
N2	2604(4.3%)	8620(20.4%)	784(4.3%)	38(11.5%)
N3	1381(2.3%)	949(2.3%)	432(2.4%)	31(9.3%)
7th AJCC M stage				
M0	60263(99.7%)	42200(99.7%)	18063(99.6%)	305(91.9%)
M1	182(0.3%)	127(0.3%)	55(0.4%)	27(8.1%)
7 th AJCC stage				
I	32929(54.4%)	23112(54.6%)	9817(54.1%)	74(22.2%)
II	21464(35.5%)	15019(35.4%)	6445(35.5%)	201(60.5%)
III	5870(9.7%)	4069(9.6%)	1801(9.9%)	56(16.8%)

IV	182(0.3%)	127(0.3%)	55(0.3%)	1(0.3%)
ER Status				
Positive	50921(84.2%)	35625(84.1%)	15296(84.5%)	246(74.0%)
Negative	9524(15.8%)	6702(15.9%)	2822(15.5%)	86(26.0%)
PR Status				
Positive	44464(73.5%)	31139(73.5%)	13325(73.5%)	214(64.5%)
Negative	15981(26.4%)	11188(26.5%	4793(26.5%)	118(35.5%)
Her-2 Status				
Positive	8217(13.5%)	5697(13.4%)	2520(13.9%)	121(36.4%)
Negative	52228(86.5%)	36630(86.6%)	15598(86.1%)	211(63.6%)
Breast Subtype				
Luminal A	45677(75.5%)	32003(75.6%)	13674(75.4%)	140(42.2%)
Luminal B	5759(9.5%)	3965(9.4%)	1794(9.9%)	34(10.2%)
Her-2 enriched	2458(4.0%)	1732(4.1%)	726(4.1%)	119(35.8%)
Triple negative	6551 (10.8%)	4627(10.9%)	1924(10.6%)	39(11.8%)
Vital status				
Alive	54354(89.9%)	38066(89.9%)	16288(89.8%)	290(87.4%)
Dead	6091 (10.1%)	4261(10.1%)	1830(10.2%)	42(12.6%)
Median follow-up time(Months, 25th−75th percentile)	48.0(34.0-65.0)	49.0(34.0-65.0)	48.0(34.0-64.0)	71.0(64.3-83.0)

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

HER2, Human epidermal growth factor 2-neu.

Table 2. Univariate and multivariate analyses of overall survival in the training corhort

Variable	OS			
	Univirate		Multivirate	
	HR(95%CI)	P Value	HR(95%CI)	P Value
Age (Years)	2.400(2.240-2.570)	0.000	2.920(2.720-3.130)	0.000
Grade	1.570(1.510-1.640)	0.000	1.250(1.190-1.310)	0.000
7 th AJCC stage	2.010(1.940-2.090)	0.000	2.020(1.940-2.110)	0.000
ER status	0.440(0.420-0.470)	0.000	0.780(0.630-0.960)	0.019
PR Status	0.510(0.480-0.540)	0.000	0.780(0.710-0.860)	0.000
Her-2 Status	0.890(0.810-0.980)	0.013	0.650(0.590-0.710)	0.000
Breast Subtype	1.340(1.310-1.380)	0.000	1.080(1.000-1.170)	0.048

AJCC, American Joint Committee on Cancer; ER, estrogen receptor; PR, progesterone receptor; HER2, Human

epidermal growth factor 2-neu; OS, overall survival.

Figures

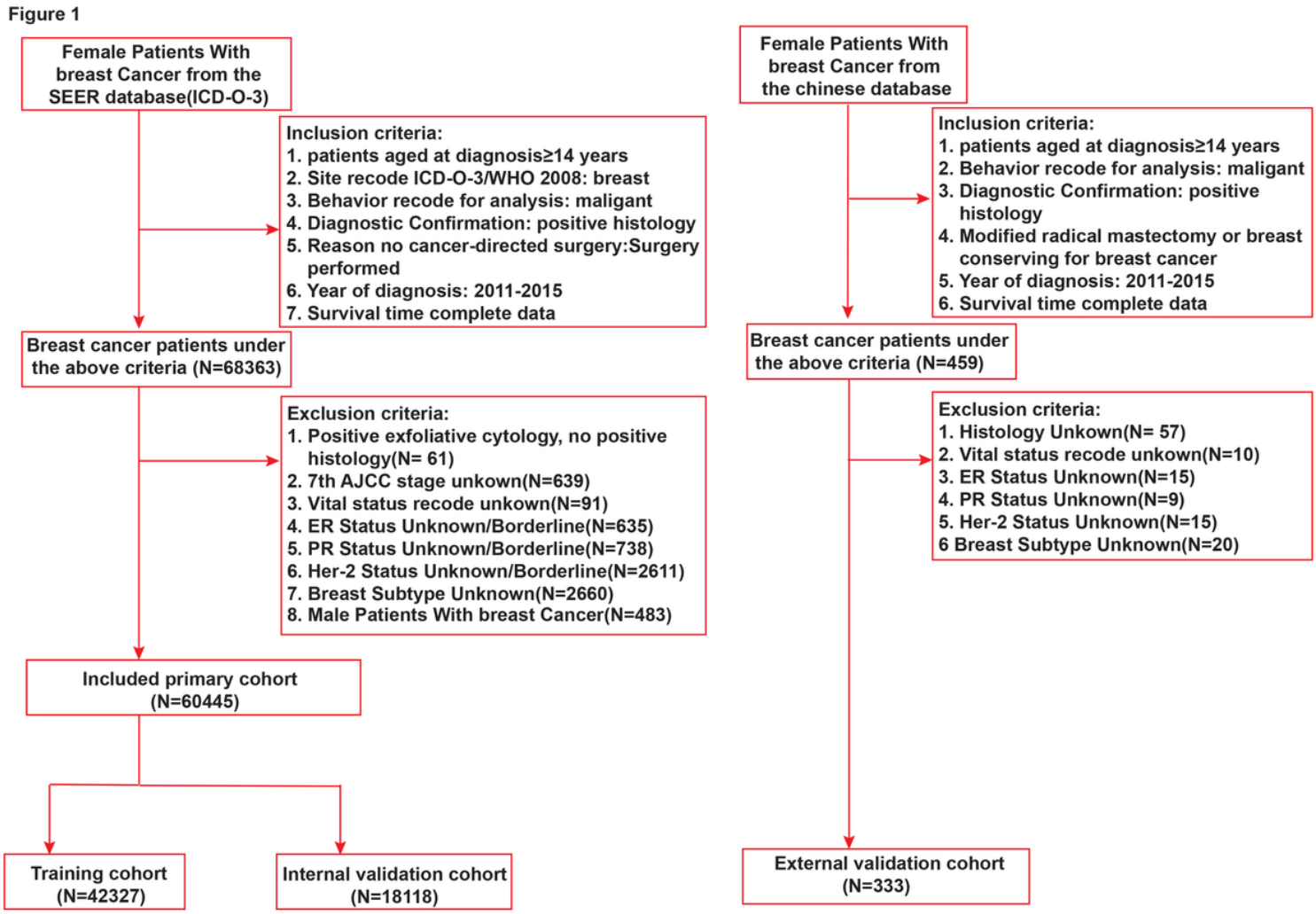

Figure 1

Flowchart of sample selection for this study. AJCC, American Joint Committee on Cancer, SEER, Surveillance, Epidemiology, and End Results, ER, estrogen receptor, PR, progesterone receptor, Her-2, Human epidermal growth factor 2-neu.

Figure 2

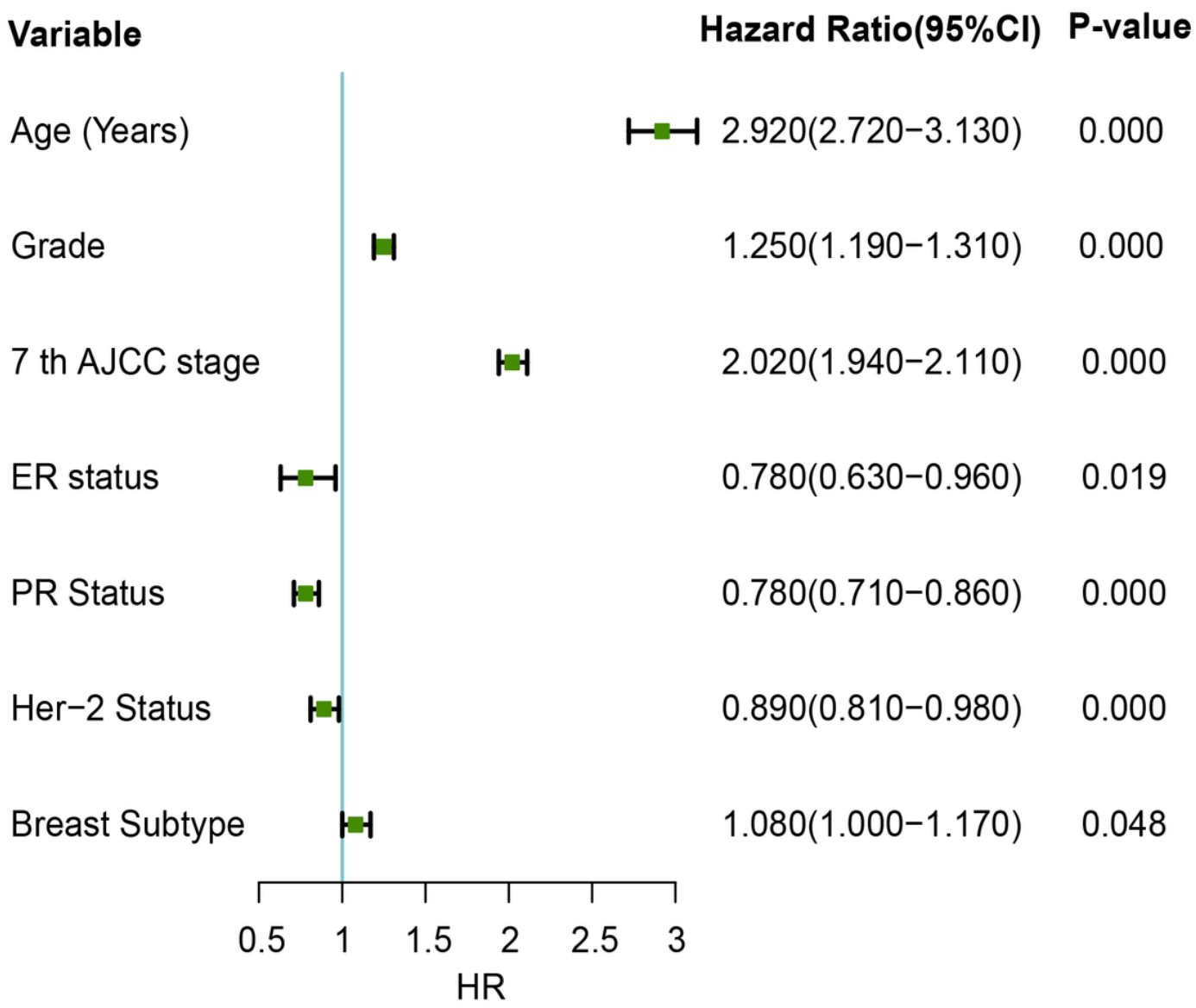

Figure 2

The effects of different prognostic factors was described by forest plot.

AJCC, American Joint Committee on Cancer, ER, estrogen receptor, PR, progesterone receptor, Her-2, Human epidermal growth factor 2-neu.

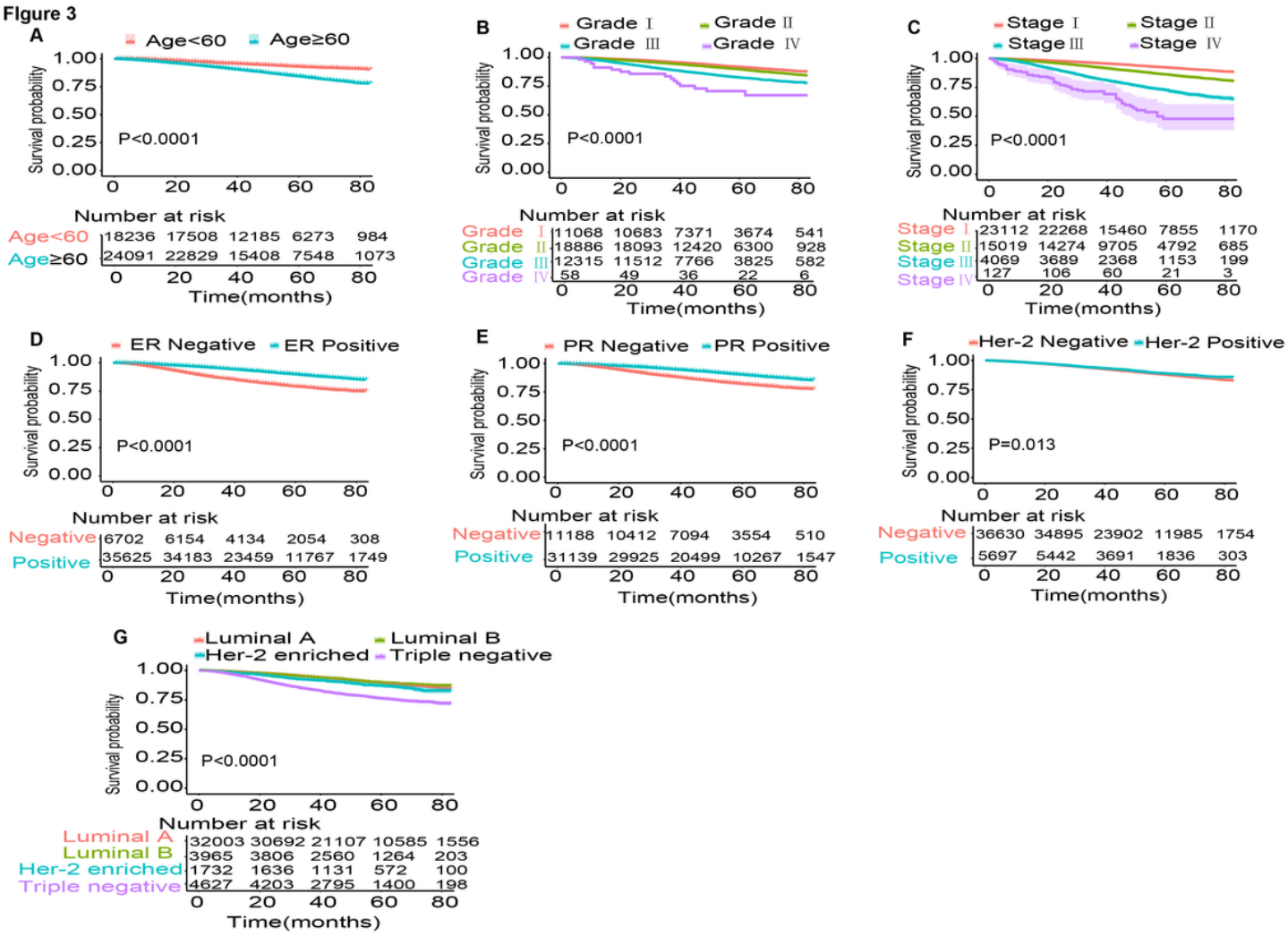

Figure 3

Kaplan-Meier OS curve of the female patients stratified by characteristics in the training corhort. (A)Age, (B)Grade, (C)7th AJCC stage, (D)ER status, (E)PR status, (F)Her-2 status, (G)Breast Subtype. AJCC, American Joint Committee on Cancer, ER, estrogen receptor, PR, progesterone receptor, Her-2, Human epidermal growth factor 2-neu, OS, overall survival.

Figure 4

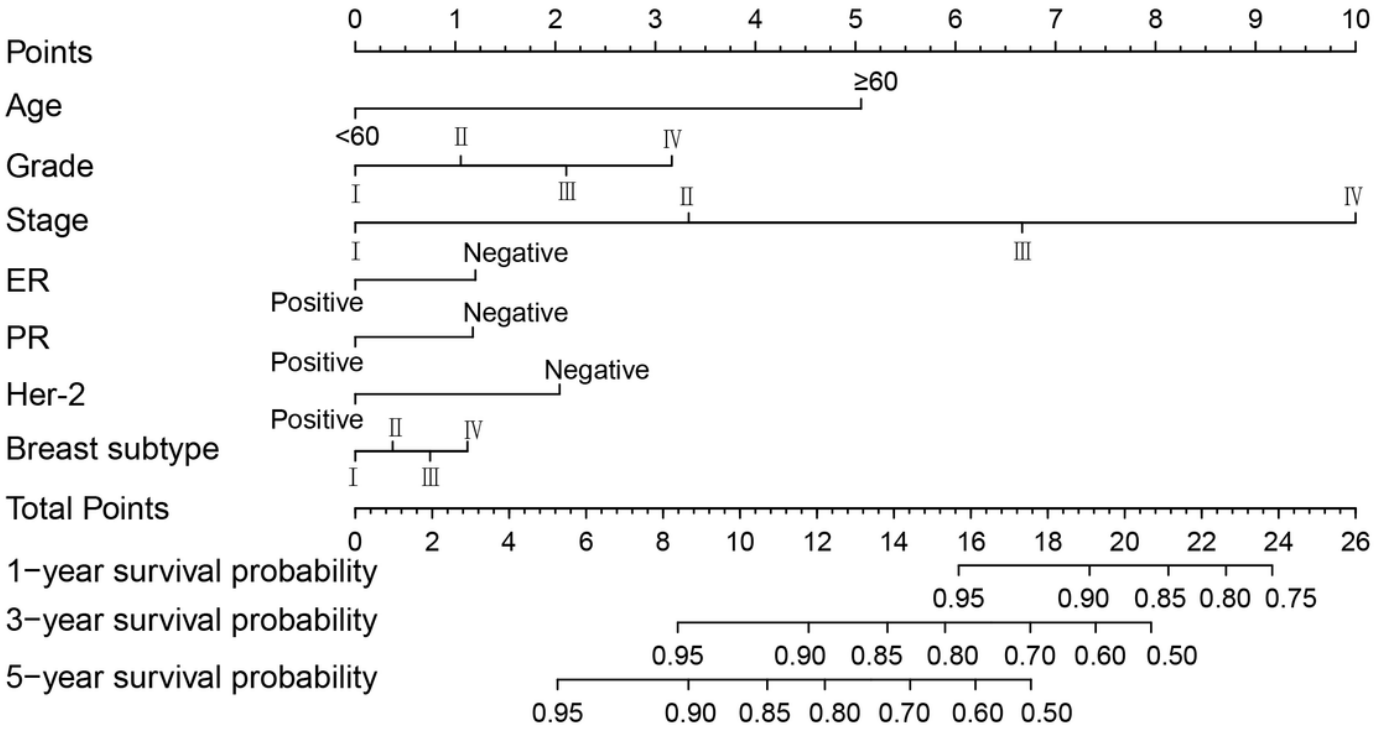

Figure 4

A nomogram predicting the 1-, 3- and 5-year OS survival of postoperative breast cancer patients in the training corhort. AJCC, American Joint Committee on Cancer, ER, estrogen receptor, PR, progesterone receptor, Her-2, Human epidermal growth factor 2-neu, OS, overall survival.

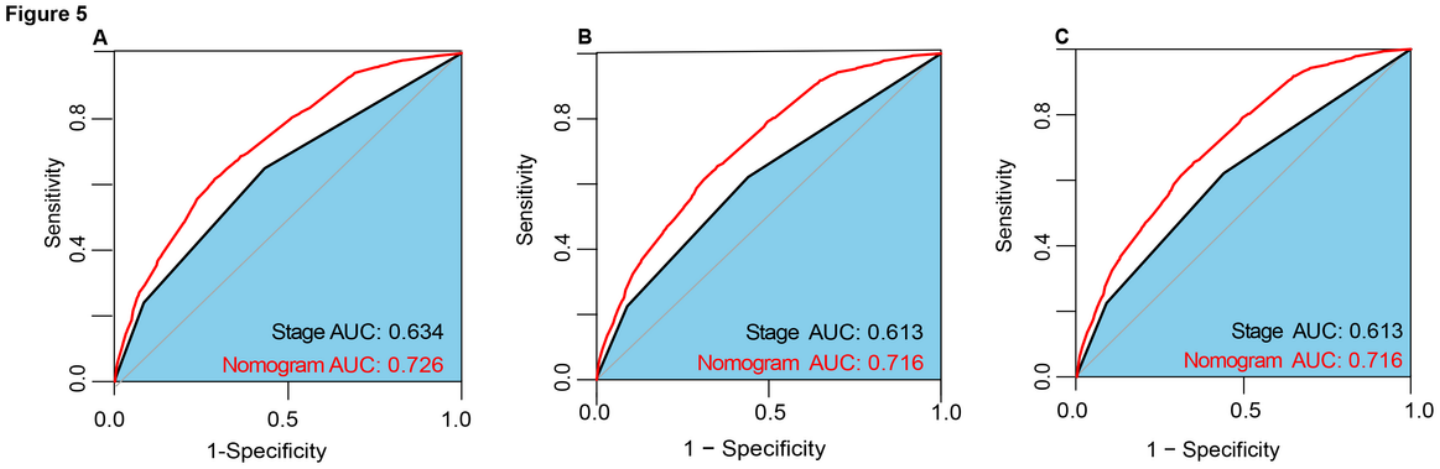

Figure 5

The ROC curves in the training cohort(A), internal validation cohort(B) and external validation cohort(C). AUC, area under the time-dependent receiver operating characteristics curve.

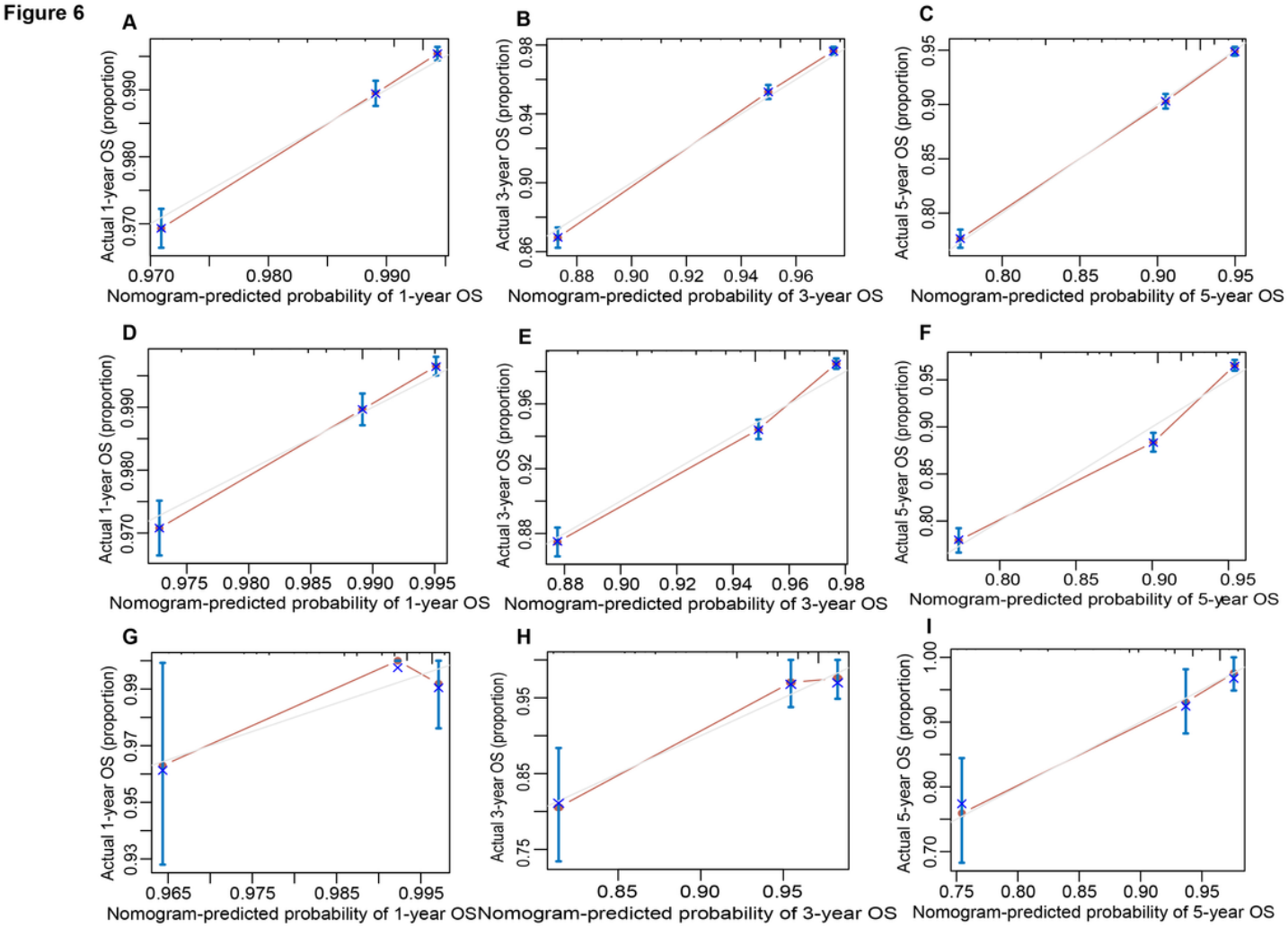

Figure 6

Calibration plot of the nomogram for predicting OS of breast cancer patients. (A) 1-year, (B) 3-year, and (C) 5-year OS according to the training cohort. (D) 1-year, (E) 3-year, and (F) 5-year OS according to the internal validation cohort. (G) 1-year, (H) 3-years, and (I) 5-years OS according to the external validation cohort. The performance is estimated by bootstrap 1,000 repetitions. The X-axis plots the nomogram-predicted survival, the Y-axis plots the actual survival. OS, overall survival.

Figure 7

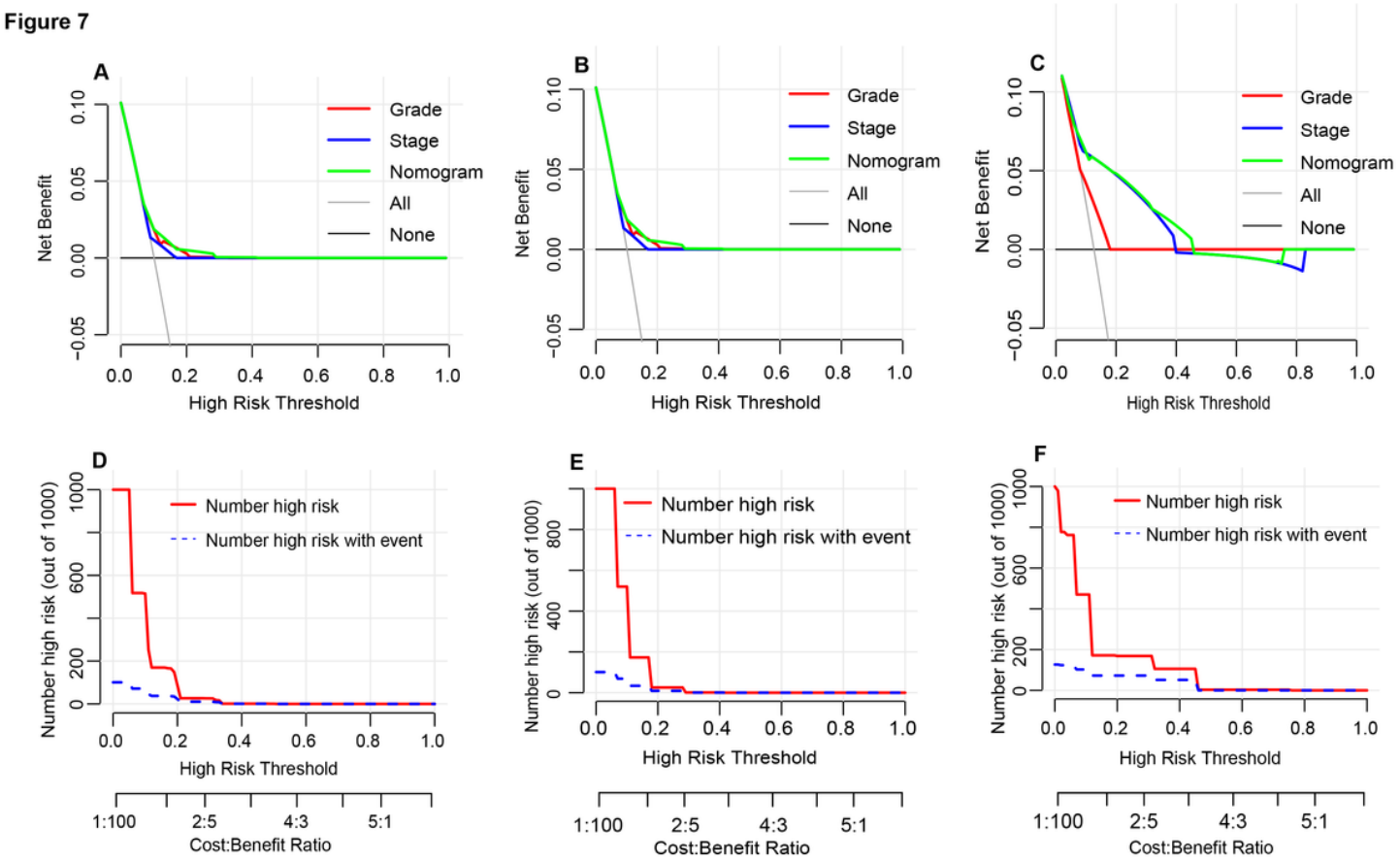

Figure 7

Decision curve analysis and clinical impact curve. Decision curve analysis for the nomogram in the training cohort (A), internal validation cohort(B) and external validation cohort(C). The horizontal solid black line represents the assumption that no patients will experience the event, and the solid gray line represents the assumption that all patients will relapse. Clinical impact curve of the nomogram for OS prediction in the training cohort (D), internal validation cohort(E) and external validation cohort(F).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GraphicalAbstractImage.pdf
- GraphicalabstractsText.doc
- Highlights.doc