

ORIGINAL RESEARCH

## Development, verification, and comparison of a risk stratification model integrating residual cancer burden to predict individual prognosis in early-stage breast cancer treated with neoadjuvant therapy

N. Hou<sup>1†</sup>, J. Wu<sup>2†</sup>, J. Xiao<sup>1†</sup>, Z. Wang<sup>1</sup>, Z. Song<sup>2</sup>, Z. Ke<sup>2</sup>, R. Wang<sup>2</sup>, M. Wei<sup>2</sup>, M. Xu<sup>2</sup>, J. Wei<sup>2</sup>, X. Qian<sup>2</sup>, X. Xu<sup>2</sup>, J. Yi<sup>1</sup>, T. Wang<sup>1</sup>, J. Zhang<sup>1</sup>, N. Li<sup>1</sup>, J. Fan<sup>1</sup>, G. Hou<sup>3</sup>, Y. Wang<sup>4</sup>, Z. Wang<sup>2\*</sup> & R. Ling<sup>1\*</sup>

Departments of <sup>1</sup>Thyroid, Breast and Vascular Surgery; <sup>2</sup>Pathology; <sup>3</sup>Urology, Xijing Hospital, Fourth Military Medical University, Xi'an; <sup>4</sup>Department of Geriatrics, Union Hospital of Tongji Medical College, Huazhong University of Science & Technology, Wuhan, PR China



Available online xxx

**Background:** A favorable model for predicting disease-free survival (DFS) and stratifying prognostic risk in breast cancer (BC) treated with neoadjuvant chemotherapy (NAC) is lacking. The aim of the current study was to formulate an excellent model specially for predicting prognosis in these patients.

**Patients and methods:** Between January 2012 and December 2015, 749 early-stage BC patients who received NAC in Xijing hospital were included. Patients were randomly assigned to a training cohort ( $n = 563$ ) and an independent cohort ( $n = 186$ ). A prognostic model was created and subsequently validated. Predictive performance and discrimination were further measured and compared with other models.

**Results:** Clinical American Joint Committee on Cancer stage, grade, estrogen receptor expression, human epidermal growth factor receptor 2 (HER2) status and treatment, Ki-67 expression, lymphovascular invasion, and residual cancer burden were identified as independent prognostic variables for BC treated with NAC. The C-index of the model consistently outperformed other available models as well as single independent factors with 0.78, 0.80, 0.75, 0.82, and 0.77 in the training cohort, independent cohort, luminal BC, HER2-positive BC, and triple-negative BC, respectively. With the optimal cut-off values (280 and 360) selected by X-tile, patients were categorized as low-risk (total points  $\leq 280$ ), moderate-risk ( $280 < \text{total points} \leq 360$ ), and high-risk (total points  $> 360$ ) groups presenting significantly different 5-year DFS of 89.9%, 56.9%, and 27.7%, respectively.

**Conclusions:** In patients with BC, the first model including residual cancer burden index was demonstrated to predict the survival of individuals with favorable performance and discrimination. Furthermore, the risk stratification generated by it could determine the risk level of recurrence in whole early-stage BC cohort and subtype-specific cohorts, help tailor personalized intensive treatment, and select comparable study cohort in clinical trials.

**Key words:** breast neoplasm, nomogram, predictive factors, prognosis, neoadjuvant chemotherapy, residual cancer burden

### INTRODUCTION

Breast cancer (BC) is the most leading malignancy, and its mortality rate ranks second among all cancer-related deaths in women.<sup>1</sup> Neoadjuvant chemotherapy (NAC) has become

an established treatment option for locally advanced BC to reduce tumor size and to increase the breast conservation rate.<sup>2,3</sup> BC patients received NAC shows a great heterogeneity of disease with a significantly different survival,<sup>4,5</sup> rendering it critical to build up a risk stratification model.

To date, existing models such as Rouzier model,<sup>6</sup> Clinical-Pathologic Scoring System incorporating estrogen receptor (ER)-negative disease and nuclear grade 3 tumor pathology (CPS + EG scoring systems),<sup>7</sup> Colleoni model,<sup>8</sup> neoadjuvant response index (NRI),<sup>9</sup> Keam model,<sup>10</sup> Nottingham Clinico-Pathological Response Index (NPRI),<sup>11</sup> and Neo-Bioscore<sup>12</sup> have been developed, validated, and extensively used in the clinic.<sup>13-15</sup> However, almost all these studies are seemingly dated because human epidermal growth factor receptor 2 (HER2)-targeted therapies and informative pathological

\*Correspondence to: Prof. Rui Ling, Department of Thyroid, Breast and Vascular Surgery, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, PR China. Tel: +86 13909210575

E-mail: [lingrui0105@163.com](mailto:lingrui0105@163.com) (R. Ling).

\*Prof. Zhe Wang, Department of Pathology, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, PR China. Tel: +86 13709220858

E-mail: [zhwang@fmmu.edu.cn](mailto:zhwang@fmmu.edu.cn) (Z. Wang).

<sup>†</sup>These authors contributed equally.

2059-7029/© 2021 The Authors. Published by Elsevier Ltd on behalf of European Society for Medical Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

measurement were not taken into consideration in that case.<sup>6-12,16,17</sup> A nomogram created for the population in the recent clinical guidelines is therefore urgently required and comparisons of the novel model with previous models are crucial.

Basic and predictive parameters for formulating a model specific to patients with BC encompass clinical stage, histology type, grade, the expression of ER, progesterone receptor (PR), Ki-67, and HER2, residual tumor burden [including pathological complete remission (pCR) status, pathological AJCC stage], and lymphovascular invasion (LVI).<sup>6-22</sup> Among these indicators, residual tumor burden appears especially important not only because it shows a negative correlation with prognosis but also its cancer burden ranges from single cancer cell to large diameters of tumor in these patients, which, in part, could lead to a great degree of heterogeneity in BC population after NAC. Several randomized studies and meta-analysis demonstrated that patients achieving pCR during the NAC have longer disease-free survival (DFS) and overall survival (OS) than those with residual cancer.<sup>23-25</sup> It is for this reason that almost all prior studies have only focused on patients with pCR or residual cancer. However, since the measurement criteria of residual cancer burden index (RCB) were proposed by Symmans et al. in 2007,<sup>26</sup> RCB score and class were highly sought-after in many studies.<sup>4,27-31</sup> RCB is based on the measurement of histopathological indicators such as number of positive nodes, diameter of the largest metastatic node, and the size and percent cellularity of the primary tumor bed ([www.mdanderson.org/breastcancer\\_RCB](http://www.mdanderson.org/breastcancer_RCB)). Recent findings showed that no prognostic difference for OS or relapse-free survival was observed between pCR and low residual cancer burden class (I),<sup>27,28,31</sup> which provides rationale to establish a predictive model for all patients (including pCR and non-pCR) who received NAC. However, there is currently no available prognostic model that can be applied for all early-stage BC patients treated with NAC.

On the other hand, individualized prediction has been regarded as an important requisite for an excellent predictive model,<sup>32-34</sup> considering that relapse is a frequent cause of death in BC and substantially affects the quality of life for BC patients. More importantly, prognostic model with excellent performance of discrimination can help develop new drugs for population in the high-risk group.<sup>35</sup> Therefore, it is crucial for clinicians to construct a novel model to accurately stratify the risk level of recurrence.

In the present study, we sought to create a risk stratification model that can represent continuous prognostic risk and can be applied to predict individual prognosis and separate patients into different risk groups in BC with NAC.

## PATIENTS AND METHODS

### Study design and population

In the current study, we collected the clinical and pathological data of 876 patients who were diagnosed with BC and received NAC at Xijing Hospital in 2012-2015. The following data were selected as study variables: age at

diagnosis, family history of BC, menopausal status, clinical AJCC stage, laterality, histology, pathological information, and treatment data. Family history of BC refers to a patient's first- or second-degree relatives with BC.

The inclusion criteria were listed as follows: (i) female BC was diagnosed by positive histology; (ii) all patients were given NAC. The exclusion criteria were as follows: (i) bilateral BC; (ii) metastatic BC; (iii) information on ER was unavailable; (iv) not received mastectomy; (v) incomplete follow-up; (vi) patients with unavailable pathology slides.

Finally, 749 patients were enrolled into the study. The eligible population was randomly assigned in a 3 : 1 ratio to a training cohort and an independent cohort by computer-generated randomized number. The flowchart of the study is shown in Figure 1A.

The clinical stage was assessed based on the sixth edition of the AJCC BC staging system before NAC.<sup>36</sup> The detailed NAC regime and dose were described in our prior reports.<sup>37</sup>

The study was approved by the institutional review board of Xijing hospital and received an exemption of informed consent from the local ethics committee.

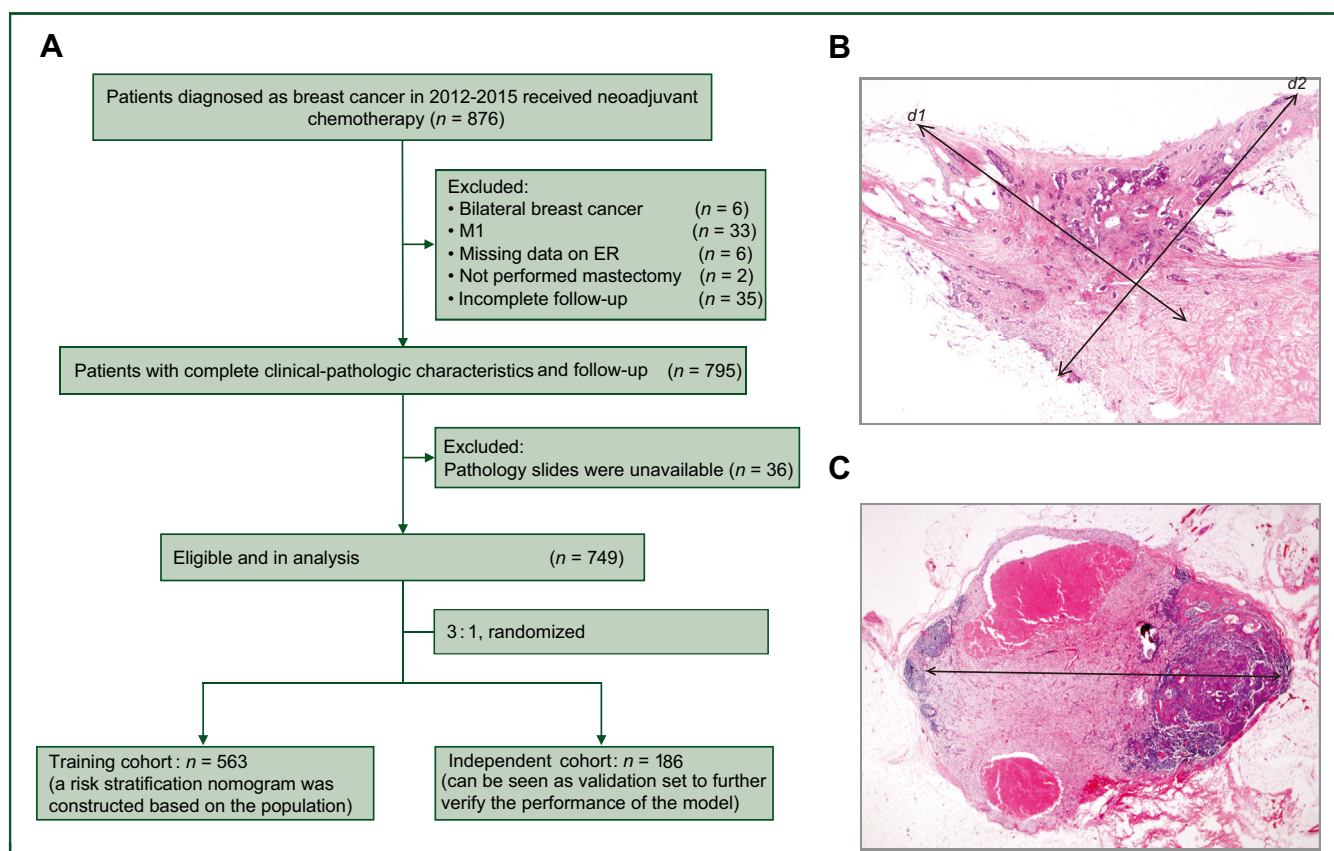
### Outcomes

The primary endpoint of our study was DFS. The secondary endpoints of this study were OS, local regional recurrence (LRR), and distant metastasis (DM). DFS was measured from the date of mastectomy to the date of disease recurrence, death, or the last follow-up. OS was defined as the time from the date of mastectomy to the date of death due to any cause or the last follow-up. LRR was defined as first cancer recurrence in the regional areas (including axillary, supraclavicular, internal mammary nodes, and chest wall). DM was identified as first recurrence beyond the LRR area as defined above. Patients with local and distant recurrence were analyzed in LRR and DM, respectively.

### Pathological evaluation of residual disease

All patients underwent a pathological evaluation at the Department of Pathology. Formalin-fixed paraffin-embedded post-NAC resection specimens were retrieved from the Pathology Department Archive. The assessment of the punctual specimens before NAC was utilized to determine the histology type of BC and grade. Considering that patients with pCR have no residual cancer burden, we carried out immunohistochemistry evaluation for ER, PR, HER2, and Ki-67 based on the pathology slides obtained before NAC. Only nuclear reactivity was considered for ER, PR, and Ki-67 expression, using continuous expression. HER2/neu positive means that the number of cells with complete membrane staining >10% of the total tumor cells or HER2 amplification determined by FISH gene detection.

Primary tumor bed, overall cancer cellularity (as percentage of area), percentage of cancer that is *in situ* disease, number of positive lymph nodes, and diameter of largest metastasis of resected lymph node after NAC were pathologically evaluated on hematoxylin–eosin-stained slides according to the criteria reported by Symmans et al.<sup>26</sup>



**Figure 1.** Flowchart of the current study design (A), parameters of residual cancer burden measurement including bidimensional diameter (d1, d2) of primary residual tumor bed (B), and the diameter of the largest nodal metastasis (C). ER, estrogen receptor.

Figure 1B and C shows the measurement of bidimensional diameters of the primary tumor bed (d1, d2) and diameter of the largest nodal metastasis. RCB was assessed and divided into RCB classes. LVI and skin involved were carefully measured based on post-NAC pathological slides. The pathology slides were measured independently by the three pathologists of the authors.

### Statistical analysis

Chi-square test or Fisher's exact test was utilized to compare the categorical variables between the training cohort and independent cohort, whereas quantitative variables were listed as median with interquartile range (IQR) and statistical comparisons were made using Student's *t*-test or non-parametric Mann–Whitney *U* test. Patients who were alive at last follow-up data (15 September 2020) or lost to follow-up were censored at the time of the last contact. The Kaplan–Meier analysis was employed to calculate the survival rate, and the log-rank test was used to compare the differences between the curves.

Variables with  $P < 0.05$  in univariable Cox analysis were incorporated into multivariable Cox analysis to generate independent prognostic factors of BC in the training cohort. A backward step-down selection identified a final model according to the Akaike information criterion.<sup>38</sup> A nomogram for predicting DFS was built up based on the

multivariable Cox results. In the current study, the bootstrap method = 200 was applied to validate the model's performance. The 3-year and 5-year receiver operating characteristic (ROC) curves and calibration curves were drawn to evaluate the predictive performance of the nomogram in the internal and independent validation.<sup>39</sup> The larger the area under the ROC curve (AUC), the higher the predictive accuracy of the nomogram. The closer the calibration curve is to the ideal curve, the more unbiased the predictive performance of the model. The time-dependent ROC curves, corresponding AUC values, and Harrell's C-indexes were utilized to measure and compare the performance of the final model, Rouzier model, CPS + EG scoring system, Colleoni model, Keam model, Neo-Bioscore, RCB, AJCC, and Ki-67<sup>39</sup> among the training cohort, independent cohort, luminal BC, HER2-positive BC, and triple-negative BC (TNBC) groups. Decision curve analysis was carried out to identify whether the nomograms could be deemed useful tools for clinical decision making by comparing the net benefits at any threshold probability.<sup>40</sup> All patients were grouped into three risk strata (low-risk, moderate-risk, and high-risk group) according to two optimal cut-offs identified by X-tile in the training cohort.<sup>41</sup>

Statistical tests were two-sided, and  $P$  values  $< 0.05$  were considered as statistically significant. All statistical analyses were conducted using R version 3.3.6 (<http://www.R-project.org/>) with packages rms, timeROC, caret, and ggDCA.

## RESULTS

### Characteristics of the training cohort and independent cohort

A total of 749 BC patients who received NAC were finally selected. Among them, 563 (75.2%) patients were included in the training cohort and 186 (24.8%) patients were included in the independent cohort (Figure 1A). Baseline characteristics between the two cohorts are presented in Supplementary Table S1, available at <https://doi.org/10.1016/j.esmooop.2021.100269>. Notably, compared with patients in the training cohort, a lower proportion of patients in the independent cohort had hormone receptor-positive (HR+) BC. Other characteristics were comparable between the two cohorts.

The median follow-up for the total population was 58 months (IQR, 50-71 months), with 3-year and 5-year DFS being 78.8% and 72.8%, respectively. The median follow-up for the training cohort and independent cohort was 58 months (IQR, 50-71 months) and 59 months (IQR, 48-71 months), respectively. Recurrence occurred in 157 (27.8%) patients in the training cohort and 53 (28.3%) patients in the independent cohort. No significant difference in survival was observed between the training cohort and independent cohort {5-year DFS 72.6% [95% confidence interval (CI), 68.8% to 76.3%] versus 73.7% (95% CI, 67.3% to 80.1%); log-rank  $P = 0.884$ } (Supplementary Figure S1, available at <https://doi.org/10.1016/j.esmooop.2021.100269>).

**Construction of a prognostic nomogram based on multi-variable prognostic analysis.** The 3-year and 5-year DFS of 563 BC patients in the training cohort were 78.5% and 72.6%, respectively. The univariable and multivariable Cox results are presented in Supplementary Table S2, available at <https://doi.org/10.1016/j.esmooop.2021.100269>. AJCC stage I, grade 1, higher expression of ER, lower expression of Ki-67, HER2 positive with trastuzumab, no LVI, and lower classes of RCB (including pCR and I) were identified as independent protective factors for BC patients treated with NAC.

A nomogram for predicting DFS was formulated by integrating all independent prognostic factors (Figure 2A).

**Internal and external validation of the nomogram.** The prognostic model-predicted DFS with 3-year and 5-year AUC was 0.82 (95% CI, 0.78-0.86) and 0.80 (95% CI, 0.76-0.84), respectively, in the internal validation, and 0.83 (95% CI, 0.76-0.90) and 0.82 (95% CI, 0.76-0.91), respectively, in the external validation of independent cohort (Figure 2B and C). Moreover, the 3-year and 5-year calibration curves of the internal validation cohort and independent cohort further presented excellent agreement between predictions and observations in the probability of 3-year and 5-year DFS (Figure 2D and E).

**Comparison of the predictive performance of the current model and other models and factors.** The performance and discrimination of the model and other models (including Rouzier model, CPS + EG scoring systems, Colleoni model,

Keam model, and Neo-Bioscore) as well as some factors (including RCB, AJCC, and Ki-67) were compared (Supplementary Table S3, available at <https://doi.org/10.1016/j.esmooop.2021.100269>). The 1-year, 2-year, 3-year, 4-year, 5-year, and 6-year AUROC values and C-indexes of the current nomogram were higher than those of other models and factors in the training cohort and independent cohort, indicating a favorable performance and discrimination (Figure 3A and B). The 3-year and 5-year decision curve analysis indicated that the net benefit of the nomogram robustly outperformed other models and single factors in the training cohort (Figure 3C and D) and independent cohort (Figure 3E and F).

### Risk stratification of the model

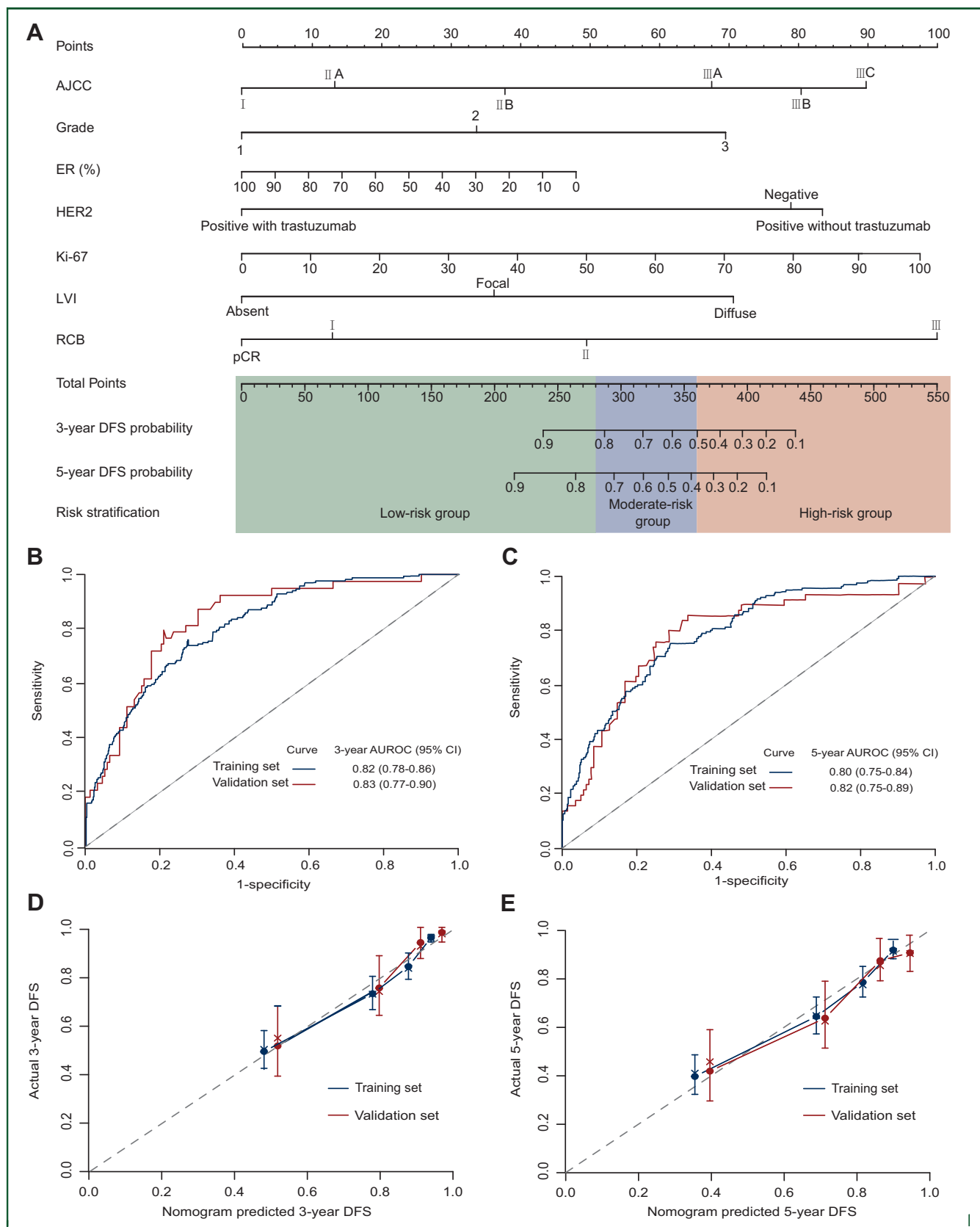
We utilized X-tile software to generate two optimal cut-offs (280 and 360, Supplementary Figure S2A-C, available at <https://doi.org/10.1016/j.esmooop.2021.100269>), which divided BC into three groups with a highly significantly different probability of recurrence (Figure 2A): low risk (total points  $\leq 280$ ,  $n = 334$  in the training cohort, and  $n = 102$  in the independent cohort), moderate risk ( $280 < \text{total points} \leq 360$ ,  $n = 167$  in the training cohort, and  $n = 65$  in the independent cohort), and high risk (total points  $> 360$ ,  $n = 62$  in the training cohort, and  $n = 19$  in the independent cohort). In the entire population, the 5-year DFS of low-risk, moderate-risk, and high-risk groups was 89.9% (95% CI, 87.0% to 92.8%), 56.9% (95% CI, 50.3% to 63.5%), and 27.7% (95% CI, 17.8% to 37.6%), respectively. With the low-risk group as reference, the hazard ratios (HRs) for moderate-risk and high-risk groups were 5.56 (95% CI, 3.91-7.91;  $P < 0.001$ ) and 13.32 (95% CI, 9.01-19.71;  $P < 0.001$ ), respectively (Figure 4A). Similar trends were also observed in the training cohort and independent cohort.

The cumulative incidence curves for recurrence and OS curves were significantly different among three groups in the training cohort and independent cohort (all with log-rank  $P < 0.001$ , Figure 4B-E). The cumulative incidence curves for LRR and DM of three groups are shown in Supplementary Figure S3, available at <https://doi.org/10.1016/j.esmooop.2021.100269>, and were significantly different (log-rank  $P < 0.001$ ).

**The performance of risk stratification model in subtype-specific BC cohorts.** The performance and discrimination of the model and other models as well as some factors were compared (Supplementary Table S4, available at <https://doi.org/10.1016/j.esmooop.2021.100269>). The AUROC values and C-indexes of the risk stratification model were higher than those of other models and factors in the luminal BC cohort, HER2-positive BC cohort, and TNBC cohort, which indicated good discrimination performance (Figure 5A-C). The 3-year and 5-year calibration curves presented good agreement between predictions and observations in the probability of 3-year and 5-year DFS among subtype-specific cohorts (Figure 5D-F).

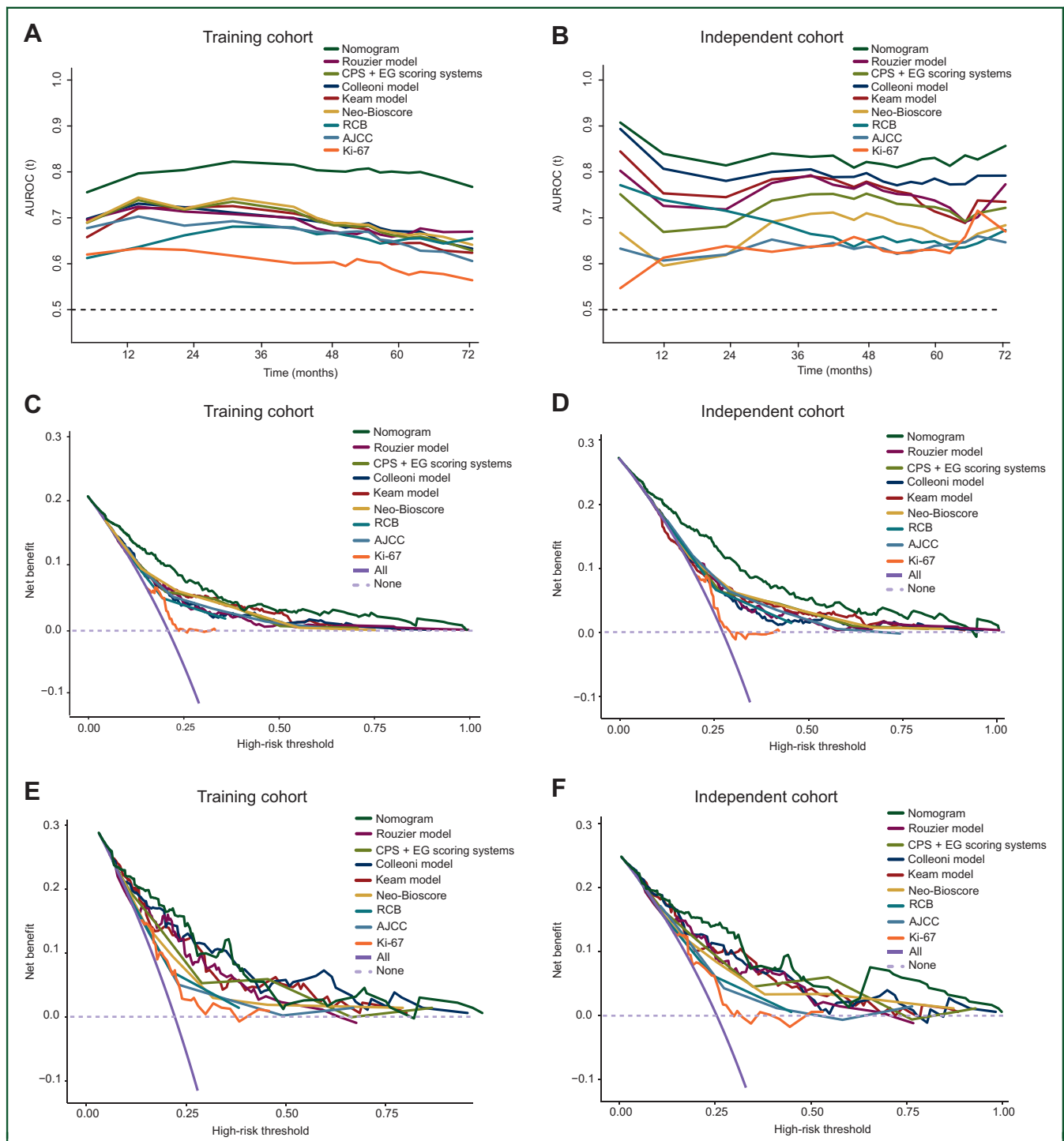
The cumulative incidence curves for recurrence and OS curves were distinctly different in the luminal BC cohort,





**Figure 2. Nomogram for individually predicting disease-free survival (DFS) in breast cancer patients treated with neoadjuvant chemotherapy (A), ROC analysis of nomogram for predicting 3-year (B) and 5-year DFS (C) in the training and independent cohorts, and calibration curves of nomogram for predicting 3-year DFS (D) and 5-year DFS (E) in the training and independent cohorts.**

AJCC, American Joint Committee on Cancer; AUROC, area under receiver operating characteristic curve; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; LVI, lymphovascular invasion; RCB, residual cancer burden.



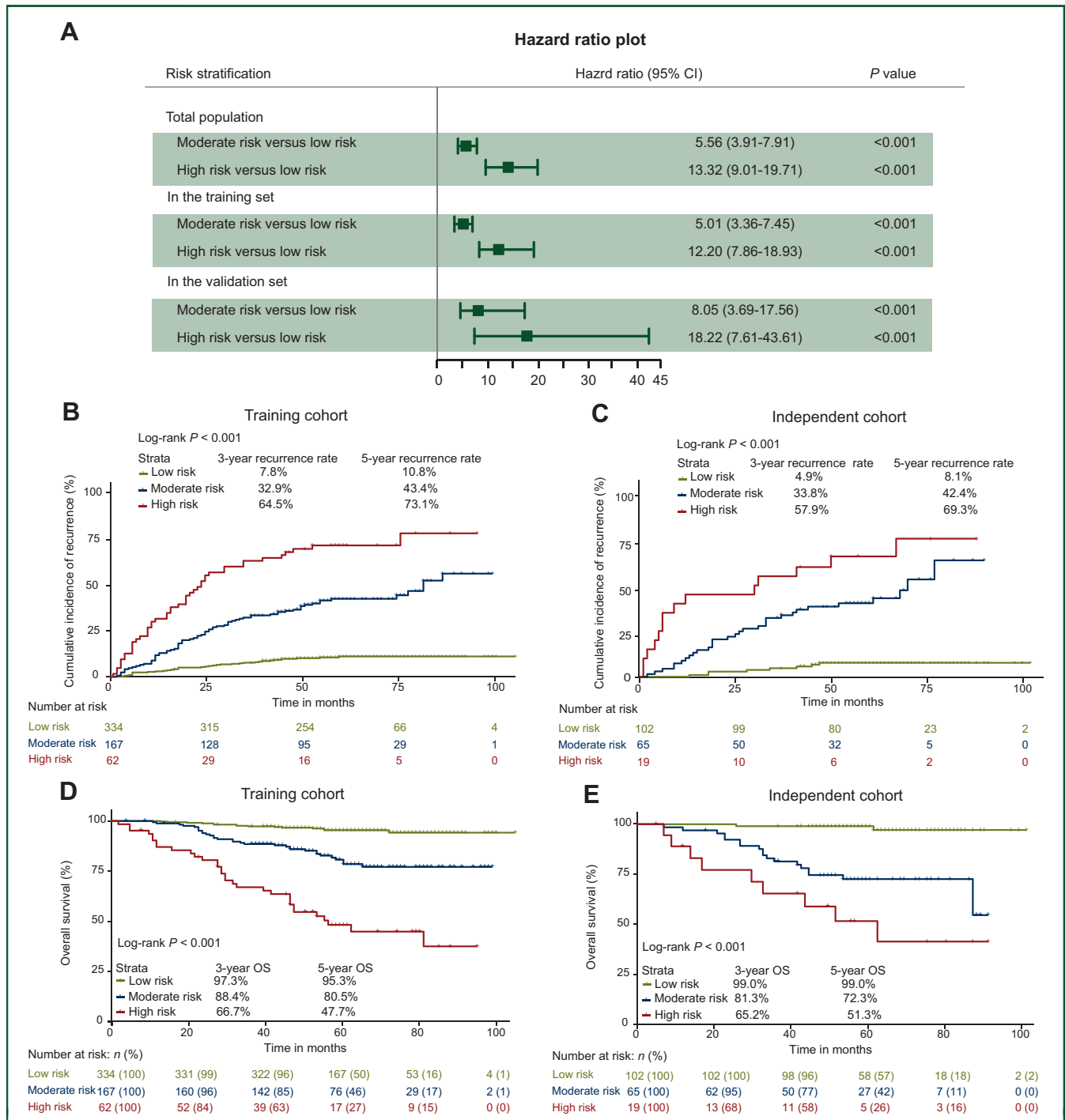
**Figure 3. Time-dependent AUC values of the current model, other available models, and single independent predictors in the training cohort (A) and independent cohort (B).**

Decision curve analysis of the model, other models, and single independent predictors for predicting the 3-year disease-free survival (DFS) (C, E) and 5-year DFS (D, F) in the training cohort and independent cohort, respectively. The threshold probability represented the 3-year (C, E) or 5-year (D, F) risk of recurrence based on each prognostic model for recommending clinical intervention. Light blue line: Assume all patients need to receive clinical intervention (all recurrence). Dotted color lines: Patients received clinical intervention if predictions exceeded a threshold, with 3-year (C, E) and 5-year (D, F) recurrence risk predictions based on different prognostic models. In general, the prognostic model with the highest net benefit at any threshold is deemed to have the highest clinical application value. AJCC, American Joint Committee on Cancer; AUC, area under receiver operating characteristic curve; RCB, residual cancer burden.

HER2-positive BC cohort, and TNBC cohort (all with log-rank  $P < 0.001$ , Figure 5G-L). There was also statistically significant difference among cumulative incidence curves for LRR and DM in the three groups (both log-rank  $P < 0.001$ , Supplementary Figure S4, available at <https://doi.org/10.1016/j.esmooop.2021.100269>).

**DISCUSSION**

Based on 749 BC patients who received NAC mainly from Northwestern China, we created the risk stratification nomogram that can estimate individualized survival with better performance and the cut-off points separated three distinct risk groups both in the overall population and the



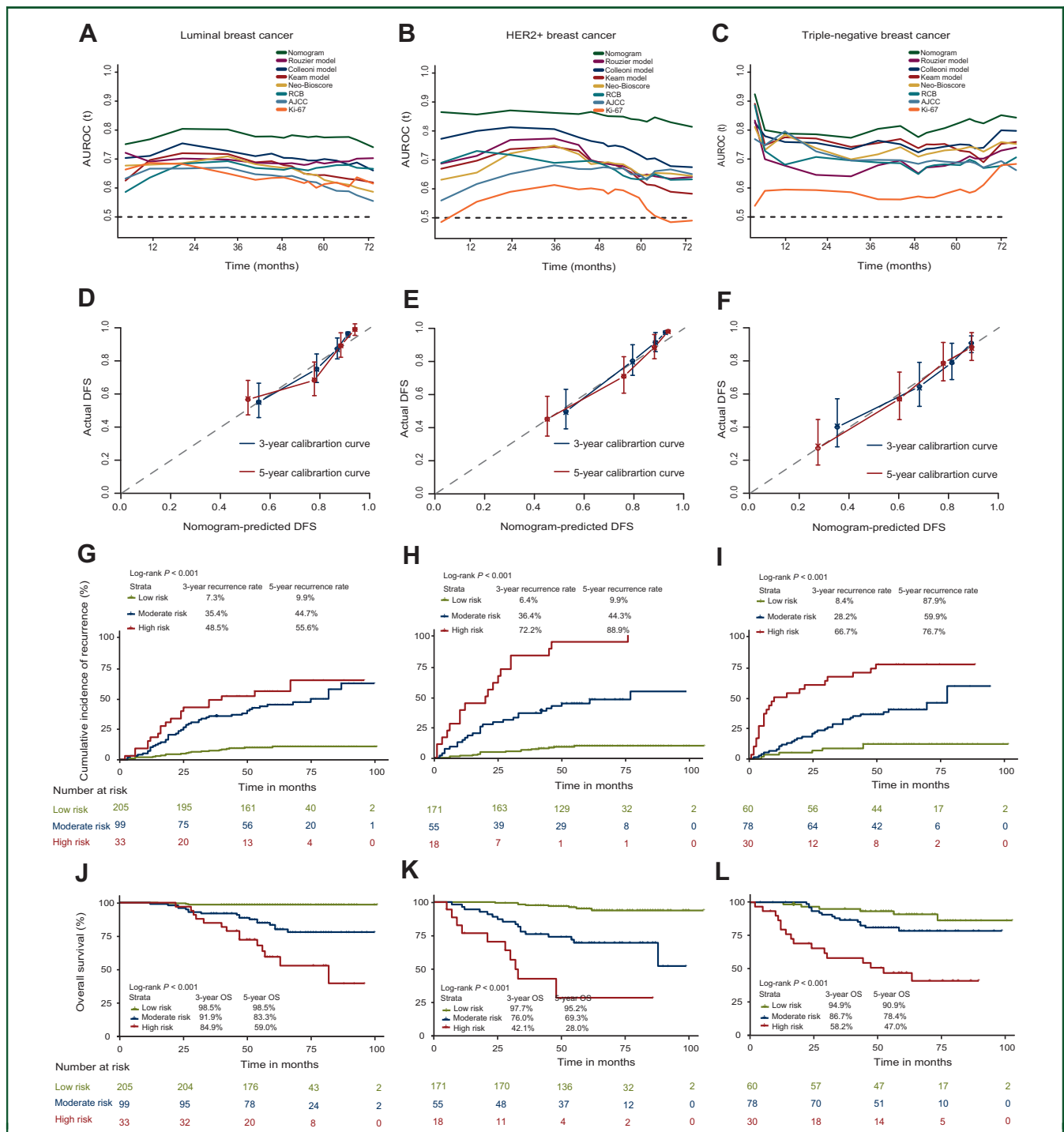
**Figure 4.** The hazard ratios of three risk stratum in the entire population, training cohort, and independent cohort (A); the cumulative incidence of recurrence curves of three risk stratum in the training cohort (B) and independent cohort (C); and overall survival curves of three risk stratum in the training cohort (D) and independent cohort (E).

CI, confidence interval; OS, overall survival.

three subtype-specific BC cohorts. The superiorities and innovative aspects lie in: (i) establishing the first model to stratify recurrence risk for BC with NAC; (ii) validating the excellent performance of the risk stratification model in the training cohort, independent cohort, and subtype-specific BC cohorts; (iii) incorporating RCB into predictive nomogram for the first time; (iv) predictive model can be utilized to predict DFS for all early-stage BC treated with NAC; and

(v) performing a continuous rather than categorized model to predict individual survival and the risk stratification can be used to select comparable population in trial design.

The 5-year DFS rates of 72.6% and 73.7% in the training and independent cohorts of all early-stage BC, respectively, were significantly higher than the 59% and 50%, respectively, noted by Colleoni et al. in 2009.<sup>8</sup> More importantly, the 5-year DFS rates of 70.1% and 71.6% in the training and



**Figure 5.** Time-dependent AUC values of the current model, other available models, and single independent predictors in subtype-specific BC cohorts (A-C); the calibration curves of nomogram for predicting 3-year DFS and 5-year DFS in subtype-specific BC cohorts (D-F); the cumulative incidence of recurrence curves of three risk stratum in subtype-specific BC cohorts (G-I); and overall survival curves of three risk stratum in subtype-specific BC cohorts (J-L). AUC, area under receiver operating characteristic curve; BC, breast cancer; DFS, disease-free survival.

independent cohorts of non-pCR patients (Supplementary Figure S5, available at <https://doi.org/10.1016/j.esmooop.2021.100269>), respectively, were obviously higher than those reported in the study by Colleoni et al., mainly because of the continued progress in therapeutic guidelines and fast improvement in drug of NAC in recent years and inclusion of patients with pCR in our study. The German BC Group reported that the 5-year DFS rate varied from 60% to

85% in 2018,<sup>42</sup> which was consistent with the present study. Patients within the low-risk group had apparently better prognosis, with an 8-year DFS of 89.4%, compared with a median DFS of 23.67 months and 77.80 months in the high-risk group and moderate-risk group, respectively (Supplementary Figure S6, available at <https://doi.org/10.1016/j.esmooop.2021.100269>). This further confirms that the survival of BC patients who received NAC varies greatly,



substantiating the significant implication of individualized estimation and risk stratification in clinical trial design and practice.

AJCC stage, grade, ER expression, HER2 status and treatment, Ki-67 expression, and LVI have been determined as independent predictors in BC patients who received NAC.<sup>6-22</sup> T stage and N stage, which indicate primary and metastatic lymph node burden, respectively, have been seen as predictors in several studies.<sup>15,18</sup> The model we constructed incorporated AJCC stage, which takes together these two factors. Specially, ER expression was utilized as a continuous variable similar to several pivotal studies and LVI was classified as no, focal, and diffused by experienced pathologists,<sup>8,43</sup> which might be more superior than two classifications. Furthermore, several studies consistently confirmed that RCB is the important prognostic indicator and outperforms other predictors. The robust predictor was included in the present study, which partly contributes to the favorable performance of the model. Additionally, it seems a bit paradoxical that post-mastectomy radiotherapy (PMRT) appears as an adverse indicator for prognosis in univariable regression analysis, yet shows no significant difference in the multivariable analysis. This is likely due to the fact that patients with PMRT have a higher proportion of high-risk group than that of non-PMRT group (15.4% versus 5.9% of high-risk group in the PMRT cohort and non-PMRT cohort, respectively,  $P < 0.001$ ). The clinical guideline of PMRT in BC also strongly substantiates this point.<sup>44</sup>

The C-index values of the novel model for predicting prognosis are superior to those previously reported in other models (0.73 and 0.67 of training and test cohorts in Colleoni model, respectively; 0.71 and 0.72 of training and test cohorts in Rouzier model, respectively; 0.78 of training in Keam model with only 2 years of follow-up). The possible explanations are that we considered RCB as a predictor and identified rational criteria for the target population in the current study.<sup>45</sup> Recently, Laas et al. reported that Neo-Bioscore had better performance in the overall population compared with RCB, but that RCB showed better performance in subtype-specific groups, especially for luminal BC and TNBC.<sup>16</sup> However, a study conducted by Dana-Farber Cancer Institute showed that RCB is superior to the Neo-Bioscore for stratifying patients into different survival outcomes.<sup>17</sup> Intriguingly, the AUROC values of five models applied to the patients in our center were inferior to those of the model. This is likely because of the distribution difference of heterogeneity in different races, the discrepancy of NAC therapeutic guidelines in different countries, and the predictive improvement of our model after including RCB and LVI.<sup>16,17</sup>

The optimal target population for NAC has been explored and determined in several randomized clinical trials, but verification regarding its prognostic difference is deficient. This may produce unrecognized confounding which might potentially affect the results of these studies, therefore a risk stratification model for discriminating and diminishing heterogeneity is urgently needed.<sup>46</sup> In particular, as is presented in [Figure 5](#) and [Supplementary Figures S3, S4, and](#)

[S6](#), available at <https://doi.org/10.1016/j.esmoop.2021.100269>, the model can separate BC patients who received NAC into different risk stratum. Especially, median DFS of the high-risk group for BC patients who received NAC is  $< 2$  years, which would severely affect the quality of life of patients in this cohort. Fortunately, the model could help clinicians to select potential high-risk cohort of recurrence and conduct clinical trials to decrease the incidence of relapse.

There are several limitations in the study: (i) The study was limited by its selective bias in retrospective study, yet this shortcoming has been diminished in identifying population based on strict criteria. (ii) Several studies have identified tumor-infiltrating lymphocytes on residual disease after NAC as an independent prognostic predictor,<sup>27,40</sup> but tumor-infiltrating lymphocytes on residual disease were not evaluated in our study. (iii) Fine-needle aspiration for metastatic axillary nodes before NAC and/or evaluation of fibrosis status in resected specimens were not routinely carried out. We were unable to compare the performance of the risk stratification nomogram with other models (e.g. NRI and NPRI), while the model was compared with most predictive models using the data in our center. (iv) Although our team considered HER2 status and treatment, not all HER2+ BC patients received HER2-targeted therapy. In fact, there do exist cases, where a predictive prognostic estimation of an HER2+ without trastuzumab BC patient is required, particularly in the developing countries. The risk stratification model we developed can also be applied in this population. (v) The newest National Comprehensive Cancer Network guidelines recommend six to eight cycles of NAC in BC patients (<https://www.nccn.org/patients/guidelines/cancers.aspx>). However, the study population received a median of four cycles of NAC (range, two to eight cycles) according to the Chinese Anti-Cancer Association guidelines for the treatment of BC in 2012-2015. Shorter cycle of NAC compared to the current guidelines may have an effect on the precise assessment of RCB. Given these drawbacks, the application of risk stratification nomogram needs to be further verified.

In conclusion, deriving from a large sample of 749 patients, the prognostic model established is the first risk stratification nomogram for predicting individual prognosis. With excellent performance and discriminative ability, the model can divide patients treated with NAC into three stratum with obviously different survival both in overall cohort and subtype-specific groups. Therefore, the risk stratification nomogram might be useful for estimating potential high-risk population of BC treated with NAC and identifying comparable candidates in clinical trials. Further verification in patients of different races remains urgently needed.

## ACKNOWLEDGEMENTS

The authors thank Prof. Jie-Lai Xia, Department of Health Statistics, Fourth Military Medical University, Xi'an, China, for his valuable statistical support.

## FUNDING

This work was supported in part by [grant number 81572917] to RL and [grant number 81902677] to ZW (Department of Thyroid, Breast and Vascular Surgery) from the National Natural Science Foundation of China.

## DISCLOSURE

The authors have declared no conflicts of interest.

## DATA SHARING

Data are available upon reasonable request. The data that support the findings of this study are available on request from the corresponding author, RL.

## REFERENCES

- DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(6):438-451.
- Mougalian SS, Soulos PR, Killelea BK, et al. Use of neoadjuvant chemotherapy for patients with stage I to III breast cancer in the United States. *Cancer*. 2015;121:2544-2552.
- Bardia A, Baselga J. Neoadjuvant therapy as a platform for drug development and approval in breast cancer. *Clin Cancer Res*. 2013;19:6360-6370.
- Campbell JL, Yau C, Krass P, et al. Comparison of residual cancer burden, American Joint Committee on Cancer staging and pathologic complete response in breast cancer after neoadjuvant chemotherapy: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). *Breast Cancer Res Treat*. 2017;165:181-191.
- Miglietta F, Dieci MV, Tsvetkova V, et al. Validation of residual proliferative cancer burden as a predictor of long-term outcome following neoadjuvant chemotherapy in patients with hormone receptor-positive/human epidermal growth receptor 2-negative breast cancer. *Oncologist*. 2020;25(9):e1355-e1362.
- Rouzier R, Pusztai L, Delaloue S, et al. Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer. *J Clin Oncol*. 2005;23:8331-8339.
- Jeruss JS, Mittendorf EA, Tucker SL, et al. Combined use of clinical and pathologic staging variables to define outcomes for breast cancer patients treated with neoadjuvant therapy. *J Clin Oncol*. 2008;26:246-252.
- Colleoni M, Bagnardi V, Rotmensz N, et al. A risk score to predict disease-free survival in patients not achieving a pathological complete remission after preoperative chemotherapy for breast cancer. *Ann Oncol*. 2009;20:1178-1184.
- Rodenhuis S, Mandjes IAM, Wesseling J, et al. A simple system for grading the response of breast cancer to neoadjuvant chemotherapy. *Ann Oncol*. 2010;21:481-487.
- Keam B, Im S, Park S, et al. Nomogram predicting clinical outcomes in breast cancer patients treated with neoadjuvant chemotherapy. *J Cancer Res Clin Oncol*. 2011;137:1301-1308.
- Abdel-Fatah TM, Ball G, Lee AHS, et al. Nottingham clinico-pathological response index (NPRI) after neoadjuvant chemotherapy (Neo-ACT) accurately predicts clinical outcome in locally advanced breast cancer. *Clin Cancer Res*. 2015;21:1052-1062.
- Mittendorf EA, Vila J, Tucker SL, et al. The neo-bioscore update for staging breast cancer treated with neoadjuvant chemotherapy. *JAMA Oncol*. 2016;2:929-936.
- Xu L, Duan X, Zhou B, et al. Validation of the CPS+EG and Neo-Bioscore staging systems after preoperative systemic therapy for breast cancer in a single center in China. *Breast*. 2018;40:29-37.
- Mittendorf EA, Jeruss JS, Tucker SL, et al. Validation of a novel staging system for disease-specific survival in patients with breast cancer treated with neoadjuvant chemotherapy. *J Clin Oncol*. 2011;29:1956-1962.
- Tanioka M, Shimizu C, Yonemori K, et al. Predictors of recurrence in breast cancer patients with a pathologic complete response after neoadjuvant chemotherapy. *Br J Cancer*. 2010;103:297-302.
- Laas E, Labrosse J, Hamy AS, et al. Determination of breast cancer prognosis after neoadjuvant chemotherapy: comparison of Residual Cancer Burden (RCB) and Neo-Bioscore. *Br J Cancer*. 2021;124(8):1421-1427.
- Kantor O, Laws A, Pastorello RG, et al. Comparison of breast cancer staging systems after neoadjuvant chemotherapy. *Ann Surg Oncol*. 2021. <https://doi.org/10.1245/s10434-021-09951-7>.
- Guarneri V, Piacentini F, Ficarra G, et al. A prognostic model based on nodal status and Ki-67 predicts the risk of recurrence and death in breast cancer patients with residual disease after preoperative chemotherapy. *Ann Oncol*. 2009;20:1193-1198.
- Hamy A, Lam G, Laas E, et al. Lymphovascular invasion after neoadjuvant chemotherapy is strongly associated with poor prognosis in breast carcinoma. *Breast Cancer Res Treat*. 2018;169:295-304.
- Klauber-DeMore N, Ollila DW, Moore DT, et al. Size of residual lymph node metastasis after neoadjuvant chemotherapy in locally advanced breast cancer patients is prognostic. *Ann Surg Oncol*. 2006;13(5):685-691.
- Fasching PA, Heusinger K, Haeberle L, et al. Ki67, chemotherapy response, and prognosis in breast cancer patients receiving neoadjuvant treatment. *BMC Cancer*. 2011;11:486.
- Buchholz TA, Tucker SL, Masullo L, et al. Predictors of local-regional recurrence after neoadjuvant chemotherapy and mastectomy without radiation. *J Clin Oncol*. 2002;20:17-23.
- Guarneri V, Broglio K, Kau SW, et al. Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors. *J Clin Oncol*. 2006;24:1037-1044.
- Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014;384:164-172.
- Broglio KR, Quintana M, Foster M, et al. Association of pathologic complete response to neoadjuvant therapy in HER2-positive breast cancer with long-term outcomes. *JAMA Oncol*. 2016;2:751-760.
- Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J Clin Oncol*. 2007;25:4414-4422.
- Symmans WF, Wei C, Gould R, et al. Long-term prognostic risk after neoadjuvant chemotherapy associated with residual cancer burden and breast cancer subtype. *J Clin Oncol*. 2017;35:1049-1060.
- Choi M, Park YH, Ahn JS, et al. Assessment of pathologic response and long-term outcome in locally advanced breast cancers after neoadjuvant chemotherapy: comparison of pathologic classification systems. *Breast Cancer Res Treat*. 2016;160:475-489.
- Luen SJ, Salgado R, Dieci MV, et al. Prognostic implications of residual disease tumor-infiltrating lymphocytes and residual cancer burden in triple-negative breast cancer patients after neoadjuvant chemotherapy. *Ann Oncol*. 2019;30:236-242.
- Steenbruggen TG, van Seijen M, Janssen LM, et al. Prognostic value of residual disease after neoadjuvant therapy in HER2-positive breast cancer evaluated by residual cancer burden, neoadjuvant response index, and Neo-Bioscore. *Clin Cancer Res*. 2019;25:4985-4992.
- Provenzano E, Bossuyt V, Viale G, et al. Standardization of pathologic evaluation and reporting of postneoadjuvant specimens in clinical trials of breast cancer: recommendations from an international working group. *Mod Pathol*. 2015;28:1185-1201.
- Li J, Lin Y, Wang Y, et al. Prognostic nomogram based on the metastatic lymph node ratio for gastric neuroendocrine tumour: SEER database analysis. *ESMO Open*. 2020;5(2):e000632.
- Lv Y, Wang Z, Li K, et al. Risk stratification based on CLIF consortium acute decompensation score in patients with Child-Pugh B cirrhosis and acute variceal bleeding. *Hepatology*. 2021;73(4):1478-1493.
- Kim J, Hong JY, Kim ST, et al. Clinical scoring system for the prediction of survival of patients with advanced gastric cancer. *ESMO Open*. 2020;5(2):e000670.

35. Prowell TM, Beaver JA, Pazdur R. Residual disease after neoadjuvant therapy—developing drugs for high-risk early breast cancer. *N Engl J Med*. 2019;380:612-615.
36. Carey LA, Metzger R, Dees EC, et al. American Joint Committee on Cancer tumor-node-metastasis stage after neoadjuvant chemotherapy and breast cancer outcome. *J Natl Cancer Inst*. 2005;97:1137-1142.
37. Hou N, Xiao J, Wang Z, et al. Development and validation of a nomogram for individually predicting pathologic complete remission after preoperative chemotherapy in Chinese breast cancer: a population-based study. *Clin Breast Cancer*. 2020;20(6):e682-e694.
38. Harrell FJ, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361-387.
39. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32:5381-5397.
40. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565-574.
41. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. 2004;10:7252-7259.
42. Denkert C, von Minckwitz G, Darb-Esfahani S, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol*. 2018;19:40-50.
43. Raphael J, Gandhi S, Li N, Lu F, Trudeau M. The role of quantitative estrogen receptor status in predicting tumor response at surgery in breast cancer patients treated with neoadjuvant chemotherapy. *Breast Cancer Res Treat*. 2017;164:285-294.
44. Recht A, Comen EA, Fine RE, et al. Postmastectomy radiotherapy: an American Society of Clinical Oncology, American Society for Radiation Oncology, and Society of Surgical Oncology Focused Guideline Update. *J Clin Oncol*. 2016;34:4431-4442.
45. Wang Q, Xia D, Bai W, et al. Development of a prognostic score for recommended TACE candidates with hepatocellular carcinoma: a multicentre observational study. *J Hepatol*. 2019;70:893-903.
46. de Azambuja E, Trapani D, Loibl S, et al. ESMO management and treatment adapted recommendations in the COVID-19 era: breast Cancer. *ESMO Open*. 2020;5:e000793.