A Model to Predict the Risk of Lymph Node Metastasis in Breast Cancer Based on Clinicopathological Characteristics

Wenxin Chen1
Chuan Wang2
Fangmeng Fu2
Binglin Yang1
Changming Chen3
Yingming Sun4

1Department of Breast Surgery, Affiliated Sanming First Hospital of Fujian Medical University, Sanming, Fujian Province 365001, People’s Republic of China; 2Breast Surgery Ward, Department of General Surgery, Fujian Medical University Union Hospital, Fuzhou, Fujian Province 350001, People’s Republic of China; 3Department of Pathology, Affiliated Sanming First Hospital of Fujian Medical University, Sanming, Fujian Province 365001, People’s Republic of China; 4Department of Radiation and Medical Oncology, Affiliated Sanming First Hospital of Fujian Medical University, Sanming, Fujian Province 365001, People’s Republic of China

Background: Sentinel lymph node biopsy (SLNB) and axillary lymph node dissection (ALND) may cause lymphatic and nervous system side effects in patients with breast cancer. It is imperative to develop a model to evaluate the risk of sentinel lymph node metastasis to avoid unnecessary operation.

Patients and Methods: A total of 2705 cases of female breast cancer patients enrolled in this retrospective study. We divided into the training group (SLNB group) and the validation group (ALND group) to analyze the relationship between lymph node metastasis and clinical-pathological factors. Logistic regression analysis was performed to verify the variables which involved in ALN metastasis and established a prediction model. ROC curves were employed to evaluate the predictive ability of the model.

Results: In the SLNB group, 9 variables, including pathological type, histological grade, tumor size, hormone receptor, HER-2, Ki-67, multifocality, and molecular subtypes, were related to breast cancer ALN metastasis. Clinically negative lymph nodes, favorable histologic type, tumor size <2 cm, and Ki-67 <15% were at very low risk for lymph node metastasis. The AUC of the validation group was 0.786.

Conclusion: We successfully establish a mathematics model to predict lymph node metastasis of breast cancer. Axillary surgery should be individual with preoperative clinical characteristics, especially for patients with a longer life expectancy.

Keywords: breast cancer, predict model, axillary lymph node metastasis, surgery

Introduction

Axillary lymph node status is one of the substantial prognostic factors for patients with breast cancer. It’s a gold standard to guide the adjuvant treatment decisions.1 Axillary lymph node dissection (ALND) is a common alternative to mastectomy in the treatment of metastatic breast cancer. However, ALND also brings many complications, including lymphedema, pain, numbness, infection, and limited shoulder movement.2,3 As more and more breast cancer gets screen-detected at an early stage, sentinel lymph node biopsy (SLNB) has emerged as a most predominant tool to probe axillary lymph node status for patients who have clinically negative lymph nodes. Review from previous data, only 30% of patients with clinically negative lymph nodes will develop sentinel lymph node metastasis.4 Although SLNB is a minimally invasive and well-tolerated surgical procedure, the short-term and long-term complications such as shoulder pain, upper limb edema, and arm numbness may also occur in patients after SLNB.5–7 It is
unreasonable for all the patients who have clinically negative lymph nodes to receive an SLNB.

As previously reported, the clinical-pathological predictors of axillary lymph node metastases are including: age,\(^8,9\) tumor size,\(^8-11,13,18\) primary tumor quadrant,\(^12\) multifocality,\(^19\) histologic grade,\(^8,14\) pathological type,\(^8,13,20\) receptor status,\(^8,21\) molecular subtypes.\(^21\) Notwithstanding, all the identified factors are independently correlated with axillary lymph node metastases. Pooling all of these predictors, alone or combined with others, has not been adopted for the selection of patients for whom ALND may be avoided.

Therefore, in this study, we were trying to figure out the correlation between clinical-pathological characteristics and axillary lymph node status in an extensive series of 2705 patients treated with conservative or ablative surgery for primary invasive breast cancer, determined the risk factors of ALN metastasis and establish a mathematics model to guide surgeons to perform SLNB on selective patients.

**Patients and Methods**

2705 patients who underwent surgical treatment between Jan 2008 and December 2018 at Fujian Medical University Union Hospital and Affiliated Sanming First Hospital of Fujian Medical University were enrolled in this study.

The inclusion criteria were as follows: (i) female; (ii) operable primary breast cancer underwent surgical treatment by tumorectomy or mastectomy; (iii) negative axillary lymph nodes confirmed by clinical and imaging examinations; (iv) SLN biopsy successfully performed and additional ALND performed; (v) Axillary lymph nodes were confirmed to be negative by histological or cytological biopsy before surgery; (vi) not pregnant; (vii) no neoadjuvant therapy and (viii) no previous history of axillary surgery. Patients who did not meet the above criteria were excluded.

The patients were divided into two groups according to the methods of axillary surgery: the SLNB group (training group) and the ALND group (validation group). In the SLNB group, patients with pathologically negative lymph nodes who underwent sentinel lymph node biopsy; in the ALND group, patients with pathologically positive lymph nodes who underwent SLNB and then received additional ALND.

The Ethics Committee approved the protocol of Fujian Medical University Union Hospital and Affiliated Sanming First Hospital of Fujian Medical University. This study was strictly conducted in accordance with the Declaration of Helsinki.

Data were collected on demographics materials and tumor clinical-pathological characteristics, including age, surgical procedure, clinical size, tumor location, postoperative pathological tumor size, pathological type, histological grade, clinical axillary lymph node status, estrogen receptor (E.R.) status, progesterone receptor (P.R.) status, Ki-67 index, human epidermal growth factor receptor 2 (HER-2) status, multifocality, postoperative axillary lymph node status, stage of the disease, molecular subtypes. All diagnostic slides of the tumors and lymph nodes were evaluated independently by two breast pathologists.

**Statistical Analysis**

The mean and S.D. were used to describe continuous variables. \(t\)-test was used for the comparison of continuous variables, and the chi-square test or Fisher’s exact test was applied for comparison of categorical variables. Univariate logistic regression analysis was applied to study ALN metastasis-related variables; multivariate logistic regression analysis was used to verify the variables that were independent affecting factors of ALN metastasis and establish a prediction model for breast cancer ALN metastasis. ROC curves were employed to evaluate the predictive ability of the model.

The predicted value of transfer risk was segmented in the decimal system according to the verification queue data, and the average transfer risk of each segment was calculated. To further evaluated the clinical application value of the model, we considered a certain predictive risk cutoff value in the verification cohort. We calculated the corresponding accuracy and false-negative rate of the cutoff value to evaluate the screening indicators of patients with low-risk ALN metastasis.

All data analyses were via SPSS (version 11.0, Chicago, IL, USA, and R software (version 3.1.0, Vienna, Austria). A two-tailed \(p\)-value <0.05 was considered as statistically significant.

**Results**

**Clinical-Pathologic Characteristics and Grouping of Patients**

As 658 patients did not meet the inclusion criteria, we enrolled a total of 2047 patients, of which 1010 were assigned into the training cohort, and 1037 were assigned into the validation cohort. The baseline of clinical-pathological data between the two groups did not differ significantly (\(p>0.05\)). (Table 1).
Table 1 Baseline of the Clinical-Pathological Characteristics Between of Two Cohorts

| Characteristics                  | Training | %     | Validation | %     | P-value |
|----------------------------------|----------|-------|------------|-------|---------|
| Number of patients               |          |       |            |       |         |
| Age                              |          |       |            |       |         |
|                                 | 1010     | 50.8±10.8 | 1037      | 50.0±10.2 | 0.082   |
| Tumor location                   |          |       |            |       |         |
| UOQ                              | 469      | 53.0  | 438        | 49.8  | 0.069   |
| LOQ                              | 122      | 13.8  | 138        | 15.7  |         |
| UIQ                              | 169      | 19.1  | 179        | 20.4  |         |
| LIQ                              | 61       | 6.9   | 51         | 5.8   |         |
| Central                          | 46       | 5.2   | 47         | 5.3   |         |
| Others                           | 18       | 2.0   | 26         | 3.0   |         |
| Clinical tumor size              |          |       |            |       |         |
| ≤2.0                             | 639      | 63.3  | 624        | 60.2  | 0.186   |
| >2.0                             | 371      | 36.7  | 413        | 39.8  |         |
| Pathological type                |          |       |            |       |         |
| IDC                              | 864      | 85.5  | 902        | 87.0  | 0.102   |
| ILC                              | 24       | 2.4   | 24         | 2.3   |         |
| Others**                         | 122      | 12.1  | 111        | 10.7  |         |
| Histologic grade                 |          |       |            |       |         |
| I                                | 62       | 6.1   | 88         | 8.5   | 0.078   |
| II                               | 521      | 51.6  | 572        | 55.2  |         |
| III                              | 319      | 31.6  | 299        | 28.8  |         |
| N.A.***                          | 108      | 10.7  | 78         | 7.5   |         |
| ER                               |          |       |            |       |         |
| Positive                         | 788      | 78    | 777        | 74.9  | 0.099   |
| Negative                         | 222      | 22    | 260        | 25.1  |         |
| PR                               |          |       |            |       |         |
| Positive                         | 661      | 65.4  | 681        | 65.7  | 0.915   |
| Negative                         | 349      | 34.6  | 356        | 34.3  |         |
| Ki-67                            |          |       |            |       |         |
| <15%                             | 244      | 24.2  | 298        | 28.7  | 0.234   |
| 15-30%                           | 415      | 41.1  | 356        | 34.3  |         |
| >30%                             | 351      | 34.8  | 383        | 36.9  |         |
| HER-2 receptor status            |          |       |            |       |         |
| Positive                         | 222      | 22    | 252        | 24.3  | 0.213   |
| Negative                         | 788      | 78    | 785        | 75.7  |         |
| Multifocality                    |          |       |            |       |         |
| Unifocal                         | 956      | 94.7  | 967        | 93.3  | 0.301   |
| Multifocal                       | 50       | 5.0   | 67         | 6.5   |         |
| Multicentric                     | 4        | 0.4   | 3          | 0.3   |         |
| Lymph node status                |          |       |            |       |         |
| Negative                         | 769      | 76.1  | 893        | 86.1  | 0.053   |
| Positive                         | 241      | 23.9  | 144        | 13.9  |         |

(Continued)
Table 1 (Continued).

| Characteristics | Training | % | Validation | % | P-value |
|-----------------|----------|---|------------|---|---------|
| Molecular subtype |          |   |            |   |         |
| ALN             | 788      | 78.0 | 777        | 74.9 | 0.110   |
| HER-2+          | 91       | 9.0  | 122        | 11.8 |         |
| TN              | 131      | 13   | 138        | 13.3 |         |

Notes: *In the case of multifocal tumors or unifocal tumors involving more than one quadrant, the tumor locations are classified in the following priority order: upper outer quadrant (UOQ), central, lower outer quadrant (LOQ), lower inner lower quadrant (LILQ), and upper inner quadrant (UILQ). For example, the inner and central regions are divided into the central quadrant. The central and outer quadrants involved are classified as outer, if all quadrants are included, the tumor is classified as the outer upper quadrant. Others (other) are occult breast cancer or tumors that cannot be found by manual examination.* 2Clinical tumor size assessment by preoperative ultrasound. 3The positivity of estrogen and progesterone receptors was defined as at least 10% or more immunostained cells. 4Positivity for HER-2 was defined as HER-2 gene amplification using the fluorescence in situ hybridization by the HER-2/neu probe kit, or scoring >3 using the immunohistochemistry method. 5Others: occult breast cancer or tumors that cannot be found by manual examination. 6Other: mucinous carcinoma, infiltrating carcinoma with medullary characteristics, squamous cell carcinoma, papillary carcinoma, infiltrating micropapillary carcinoma, strom carcinoma, medullary carcinoma, tubular carcinoma, etc. 666Some specific types of cancer are not histologically graded. N.A, not available.

Some Factors Were Involved in ALN Metastasis in the Training Cohort

Univariate logistic regression was performed to explore ALN metastasis-related variables. Pathological type, histologic grade, tumor size, hormone receptors, HER-2, Ki-67, multifocality, and molecular subtypes were attributed to ALN metastasis (p <0.05) (Table 2).

Five clinical-pathological predictors, including Pathological type, HER-2, Ki-67, Hormone receptors, tumor size, were involved in the multivariate logistic regression model. (Table 3).

Establishment of Mathematics Model for ALN Metastasis

According to the results of multivariate analysis, the following 5 variables were included in the prediction model of ALN metastasis: pathological type, HER-2, Ki-67, hormone receptors, and tumor size. The weights of each variable in the model corresponded to different point values. Points for the following factors were added together to determine total points, which corresponded to the linear predictors and risk predictors of ALN metastasis: pathology (LDC=0; IDC=100; others=31), HER-2 (Positive=0; Negative=43), Ki-67 (High=0; Low=5; Moderate=55), Hormone receptors (HR=0; HR+=32), size (<2.0 cm=0; >2.0 cm=71) (Figure 1). The final logistic regression model for predicting lymph node metastasis listed as follows:

\[
\ln(p/(1-p)) = -1.623 \times a2 - 1.167 \times a3 - 0.749 \times b1 + 0.939 \times c2 - 0.201 \times c3 - 0.502 \times d1 + 1.225 \times e1 - 2.537
\]

where “p” represents the risk of ALN metastasis. “a” represents Histological type (a2=IDC; a3=others); “b” represents HER-2 status (b1=HER-2+); “c” represents Ki-67 expression (c2=low; c3=moderate); “d” represents Hormone receptors (d1=HER-2-); “e” represents tumor size (e1=size >2.0 cm).

We retrospectively utilized the model upon the patients in the training cohort (n=1010). The area under curve (AUC) value was 0.725 (Figure 2), suggesting the predictive ability of this model was formidable.

The Predictive Ability of the Prediction Model in the Validation Cohort

We utilized the model on the patients in the validation cohort. According to the ROC curve, the AUC value was 0.786 (Figure 3), indicating the model performed well in the validation cohort as well. To further evaluate the clinical value of the model under different risk of ALN metastasis, we selected 10 cutoffs according to the Youden’s index. We calculated the number of patients, number of patients with ALN metastasis, sensitivity, specificity, accuracy, positive predictive value, and negative predictive value of ALN metastasis under different predicted risk values. As shown in Table 4, this model is more appropriate for predicting the negative axillary lymph nodes with the negative predictive rate >90%. Only 5 patients (2.9%) had lymph node metastasis among 172 cases with a predicted risk of <10%. However, the model presented a poorer prediction of positive lymph nodes. As a result, our model was more applicable to predict negative lymph nodes.

Discussion

Sentinel lymph node biopsy has become the standard surgical procedure to determine the status of axillary lymph node metastasis in breast cancers with clinically negative lymph nodes. Of note, SLNB may result in
many short-term or long-term complications. In this study, we successfully established a nomogram model for the prediction of breast cancer ALN metastasis. In our nomogram model, the histopathological type comes as the most weighted factor; Meanwhile, tumor size, Ki-67, HER-2 expression, and hormone receptors contribute to ALN metastasis.

Several studies demonstrated that age at diagnosis, tumor size, primary tumor quadrant, presence of multiple tumors, histologic grade, pathological type, ER/PR, HER-2 status, molecular subtypes, and other factors were related to ALN metastasis status.8,22–24 Viale et al.23 studied data from 4351 patients and found favorable historical types (medullary, cribriform, tubular, mucinous) have a significantly lower risk of axillary lymph node metastasis than ductal. Similarly, a lower frequency of axillary nodal metastasis in invasive lobular carcinoma than in invasive duct carcinoma has been

| Variables                  | Coefficient | Wald Value | P-value | OR   | OR 95% CI Lower | OR 95% CI Upper |
|----------------------------|-------------|------------|---------|------|-----------------|-----------------|
| Age                        |             |            |         |      |                 |                 |
| 35–44 VS <35               | 0.264       | 7.399      | 0.193   | 1.302| 0.598           | 2.863           |
| 45–54 VS <35               | 0.659       | 2.944      | 0.086   | 1.933| 0.91            | 4.103           |
| 55–64 VS <35               | 0.337       | 0.718      | 0.397   | 1.401| 0.642           | 3.053           |
| 65–74 VS <35               | 0.44        | 0.967      | 0.325   | 1.552| 0.646           | 3.726           |
| >75 VS <35                 | −0.022      | 0.001      | 0.973   | 0.978| 0.269           | 3.555           |
| Pathological type          |             |            |         |      |                 |                 |
| ILC VS IDC                 | −0.564      | 20.363     | <0.001  | 0.569| 0.298           | 1.087           |
| Other VS IDC               | −0.737      | 17.943     | <0.001  | 0.479| 0.34            | 0.673           |
| Other VS ILC               | −0.173      | 0.22       | 0.639   | 0.842| 0.409           | 1.731           |
| Histologic grade           |             |            |         |      |                 |                 |
| II V SI                    | 0.546       | 26.987     | <0.001  | 1.727| 1.23            | 2.425           |
| III V SI                   | 0.741       | 9.971      | 0.002   | 2.097| 1.475           | 2.983           |
| NO V SI                    | −0.169      | 0.347      | 0.556   | 0.844| 0.481           | 1.483           |
| Clinical tumor size        |             |            |         |      |                 |                 |
| >2.0 VS ≤2.0              | 0.493       | 9.896      | 0.002   | 1.637| 1.204           | 2.225           |
| Hr- VS HR+                 | −0.519      | 7.016      | 0.008   | 0.595| 0.405           | 0.874           |
| HER-2+ VS HER-2-           | −0.407      | 4.513      | 0.034   | 0.666| 0.458           | 0.969           |
| Ki-67                      |             |            |         |      |                 |                 |
| 15–30% VS <15%             | 0.551       | 16.286     | <0.001  | 1.735| 1.19            | 2.529           |
| >30% VS <15%               | −0.082      | 8.221      | 0.004   | 0.922| 0.61            | 1.392           |
| 15–30% VS >30%             | 0.322       | 0.151      | 0.698   | 1.38 | 1.132           | 1.682           |
| Tumor location             |             |            |         |      |                 |                 |
| LOQ VS UOQ                 | 0.24        | 5.449      | 0.364   | 1.132| 0.727           | 1.762           |
| UIQ VS UOQ                 | −0.382      | 0.3        | 0.584   | 1.132| 0.443           | 1.052           |
| LIQ VS UOQ                 | −0.177      | 2.996      | 0.083   | 0.683| 0.446           | 1.575           |
| Central VS UOQ             | −0.379      | 0.301      | 0.583   | 0.838| 0.321           | 1.459           |
| Other VS UOQ               | 0.341       | 0.965      | 0.326   | 0.684| 0.517           | 3.828           |
| Multifocality              |             |            |         |      |                 |                 |
| Multifocal VS Unifocal     | 0.656       | 4.011      | 0.135   | 1.928| 1.014           | 3.668           |
| Multicentric VS Unifocal   | 0.043       | 4.011      | 0.045   | 1.044| 0.108           | 10.095          |
| Molecular subtype          |             |            |         |      |                 |                 |
| HER-2+ VS LM               | −0.412      | 0.72       | 0.027   | 0.662| 0.382           | 1.148           |
| TN VS LM                   | −0.598      | 2.154      | 0.142   | 0.662| 0.336           | 0.901           |
| TN VS HER-2+               | −0.185      | 0.266      | 0.606   | 0.831| 0.411           | 1.68            |
reported in several studies. In our study, invasive ductal carcinoma indicated the most unfavorable outcome of ALN metastasis. The lowest rate of node positivity was observed in invasive lobular carcinoma.

Tumor size emerged as the most influential independent predictor of SLN metastases. This fact has been confirmed in some studies. Wada et al., have reported that larger primary tumors (>2.0cm) were significantly associated with positive non-SLNs (NSLN+). In a series of 893 cases published by Cutuli et al., the rate of lymph node metastases increased from 11% to 36% when the tumor size increased from 10 mm to 25 mm. In our multivariate model, the influence coefficient of tumor size is 1.225; the effect weight is just a little bit lower than the histological type. The figures are entirely in line with previous findings on the likelihood of axillary lymph node involvement after ALND according to tumor size.8–15,19

Ki-67 is a nuclear protein that is associated with cellular proliferation. Many studies26,28 suggested that higher Ki-67 index results were significantly associated with more frequency of ALNM in breast cancer. In several studies, patients were categorized into two categories of above and below 20% for the Ki-67 index. Our

| Variables             | Coefficient | SE    | Wald Value | P value | OR  | OR 95% CI  |
|-----------------------|-------------|-------|------------|---------|-----|------------|
|                       |             |       |            |         |     |            |
| Pathological type     |             |       |            |         |     |            |
| ILC VS LDC            | −1.623      | 0.776 | 16.711     | <0.001  | 0.197| 0.043–0.903|
| Other VS LDC          | −1.167      | 0.324 | 12.968     | <0.001  | 0.311| 0.165–0.588|
| HER-2                 |             |       |            |         |     |            |
| HER-2+ VS HER2-       | −0.749      | 0.240 | 9.761      | 0.002   | 0.473| 0.295–0.756|
| Ki-67                 |             |       |            |         |     |            |
| 15–30% VS ≤15%        | 0.939       | 0.226 | 35.650     | <0.001  | 2.558| 1.643–3.985|
| >30% VS <15%          | −0.201      | 0.269 | 17.269     | <0.001  | 0.457| 0.183–1.387|
| Hormone receptors     |             |       |            |         |     |            |
| HR- VS HR+            | −0.502      | 0.255 | 3.857      | 0.050   | 0.606| 0.367–0.999|
| Tumor size            |             |       |            |         |     |            |
| >2.0 VS ≤2.0          | 1.225       | 0.177 | 48.020     | <0.001  | 3.403| 2.407–4.812|
| Constant              | −2.537      | 0.473 | 28.727     | <0.001  | 0.079|             |

Figure 1 Nomogram for predicting the probability of axillary lymph node metastasis. There are a total of 8 rows in the nomogram. The behavioral variables are presented in rows 2 to 6, and points for each variable correspond to the scale in row 1. The points of these 5 variables were added to the total score shown in the balance in row 7, which relevant to the risk predictors of axillary lymph node metastasis in row 8.

Figure 2 Receiver operating characteristic (ROC) curve calculation for the predictive model of the training cohort (n = 1010).
study seems to indicate that moderately malignant tumors and low-grade malignant tumors are most prone to axillary lymph node metastasis.

There are many available nomogram models for predicting ALND metastasis. In 2007, MSKCC put forward a model that age, tumor size, tumor type, lymphovascular invasion, tumor location, multifocality, and estrogen and progesterone receptors were associated with SLN metastasis. 29 In 2012, Chen et al established an SCH model, and they identified age, tumor size, tumor location, tumor type, and lymphovascular invasion were independent predictors of SLN metastasis. 30

Owing to skipping metastasis and unproficiency of doctors, SLNB has a false-negative probability of about 10%. We aimed to identify which patients had a low risk of both SLND and ALND metastasis and to protect them from invasive procedures. Therefore, we only targeted the study population to patients with clinically negative axillary lymph nodes. We used the clinical and pathological data of patients undergoing sentinel lymph node biopsy to create a model of axillary lymph node metastasis. Compared with the MSKCC and SCH nomogram, our model has the same AUC with fewer variables, and the false-negative rate is less than 10% in both training and validation groups.

Still, we have some limitations to this study. It would be unusual management in the validation group, patients who were clinically and radiologically node-negative, yet had ALND. As most of our data were over ten years, the SLNB had not been fully promoted in our centers at that time. However, only very few patients received ALND without SLNB enrolled in this analysis. We reported the majority of patients in the ALND group were those who had a positive for SLNB and continued to receive an ALND. Moreover, the data only came from two centers, and the nomogram still needs to be further validated in various patient populations to demonstrate its reproducibility.

In conclusion, our nomogram could serve as an acceptable and adoptable clinical tool in preoperative evaluation, especially for those who were clinically and radiologically node-negative to avoid an SLNB procedure safely—thus improving the quality of life of the patients without adverse effect on their survival rates.

Table 4 Accuracy of the Developed Model in the Validation Cohort

| Predicted Risk | No. of Patients (%) | No. of Patients with ALN Metastasis | Sensitivity (%) | Specificity (%) | Accuracy (%) | Positive Predictive Rate (%) | Negative Predictive Rate (%) |
|----------------|---------------------|-----------------------------------|-----------------|-----------------|-------------|-----------------------------|-----------------------------|
| <0.051         | 43 (41.15)          | 1                                 | 99.31           | 4.70            | 17.84       | 14.39                       | 97.67                       |
| <0.082         | 105 (10.13)         | 3                                 | 97.92           | 11.42           | 23.43       | 15.13                       | 97.14                       |
| <0.100         | 172 (16.59)         | 5                                 | 96.53           | 18.70           | 29.51       | 16.07                       | 97.09                       |
| <0.142         | 313 (30.18)         | 12                                | 91.67           | 33.71           | 41.76       | 18.23                       | 96.17                       |
| <0.163         | 471 (45.42)         | 19                                | 86.81           | 50.62           | 55.64       | 22.08                       | 95.97                       |
| <0.205         | 503 (48.51)         | 20                                | 86.11           | 54.09           | 58.53       | 23.22                       | 96.02                       |
| <0.261         | 602 (58.05)         | 38                                | 73.61           | 63.16           | 64.61       | 24.37                       | 93.69                       |
| <0.313         | 737 (71.07)         | 50                                | 65.28           | 76.93           | 75.31       | 31.33                       | 93.22                       |
| <0.407         | 903 (87.08)         | 69                                | 52.08           | 93.39           | 87.66       | 55.97                       | 92.36                       |
| <0.487         | 920 (88.72)         | 75                                | 47.92           | 94.62           | 88.14       | 58.97                       | 91.85                       |

Figure 3 Receiver operating characteristic (ROC) curve calculation for the predictive model of the validation cohort (n = 1037).
The Statement of Consent Obtaining
The consent was obtained from the study participants before study commencement via signing the Informed Consent Form.

Author Contributions
All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis, and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work. Wenxin Chen and Yingming Sun made a language copy editing for the revised manuscript.

Funding
This study was supported by Sanming Science and Technology Project (No. 2017-5-3) to Prof. Wenxin Chen “The Sailing Fund of Fujian Medical University (No.2018QH1164)” and “The Scientific Fund of Sanming Science and Technology Bureau (N0.2018-5-1(7))” to Yingming Sun M.D.,Ph.D. Special thanks to Dr. Tejinder from the Ohio State University for his careful language copy editing.

Disclosure
The authors have declared that no competing interest exists.

References
1. Quiet CA, Ferguson DJ, Weichselbaum RR, Hellman S. Natural history of node-negative breast cancer: a study of 826 patients with long-term follow-up. J Clin Oncol. 1995;13:1144–1151. doi:10.1200/JCO.1995.13.5.1144
2. Roses DF, Brooks AD, Harris MN, Shapiro RL. Complications of level I and II axillary dissection in the treatment of carcinoma of the breast. Ann Surg. 1999;230:194–201. doi:10.1097/00002455-199908000-000009
3. Warmuth MA, Bowen G, Prosniitz LR, Chu L, Broadwater G, Peterson B. Complications of axillary lymph node dissection for carcinoma of the breast: a report based on a patient survey. Cancer. 1998;83:1362–1368. doi:10.1002/(SICI)1097-0142(19980101)83:7<1362::AID-CNCR13>3.0.CO;2-2
4. Krag DN, Anderson SJ, Julian TB, et al. Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised Phase 3 trial. Lancet Oncol. 2010;11:927–933. doi:10.1016/S1470-2241(10)70207-2
5. Del Bianco P, Zavagno G, Barelli P, et al. Morbidity comparison of sentinel lymph node biopsy versus conventional axillary lymph node dissection for breast cancer patients: results of the sentinel-GIVOM Italian randomised clinical trial. Eur J Surg Oncol. 2008;34:508–513. doi:10.1016/j.ejso.2007.05.017
6. Rönkä R, von Smitten K. One-year morbidity after sentinel node biopsy and breast surgery. Breast. 2005;14:28–36. doi:10.1016/j.brest.2004.09.010
7. Sener SF, Winchester DJ, Martz CH, et al. Lymphedema after sentinel lymphadenectomy for breast carcinoma. Cancer. 2001;92:748–752. doi:10.1002/1097-0142(20010815)92:4<748::AID-CNCR1378>3.0.CO;2-V
8. Gann PH, Colilla SA, Gapstur SM, Winchester DJ. Factors associated with axillary lymph node metastasis from breast carcinoma: descriptive and predictive analyses. Cancer. 1999;86:1511–1519.
9. Gajdos C, Tarter PI. Lymphatic invasion, tumor size, and age are independent predictors of axillary lymph node metastases in women with T1 breast cancers. Ann Surg. 1999;230:692–696. doi:10.1097/00002455-199911000-00012
10. Chua B, Ung O. Frequency and predictors of axillary lymph node metastases in invasive breast cancer. ANZ J Surg. 2001;71:723–728. doi:10.1046/j.1445-1433.2001.02266.x
11. Van Zee KJ, Manasseh DM, Bevilacqua JL, et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. Ann Surg Oncol. 2003;10:1140–1151. doi:10.1245/S000348012003.03.015
12. Zhang J, Li X, Huang R, et al. A nomogram to predict the probability of axillary lymph node metastasis in female patients with breast cancer in China: a nationwide, multicenter, 10-year epidemiological study. Oncotarget. 2017;8:35311–35325. doi:10.18632/oncotarget.13330
13. Silverstein MJ, Skinner KA. Predicting axillary nodal positivity in 2282 patients with breast carcinoma. World J Surg. 2001;25:767–772. doi:10.1007/s00268-001-0003-x
14. Barth A, Craig PH. Predictors of axillary lymph node metastases in patients with T1 breast carcinoma. Cancer. 1997;79:1918–1922. doi:10.1002/(SICI)1097-0142(19970515)79:10<1918::AID-CNCR12>3.0.CO;2-Y
15. Chadha M, Chabon AB. Predictors of axillary lymph node metastases in patients with T1 breast cancer. A multivariate analysis. Cancer. 1994;73:350–353. doi:10.1002/1097-0142(19940115)73:2<350::AID-CANCER2820703219>3.0.CO;2-5
16. Giuliano AE, Barth AM, Spivack B, Beitsch PD. Incidence and predictors of axillary metastasis in T1 carcinoma of the breast. J Am Coll Surg. 1996;183:185–189.
17. Wada N, Imoto S, Yamauchi C. Predictors of tumour involvement in remaining axillary lymph nodes of breast cancer patients with positive sentinel lymph node. Eur J Surg Oncol. 2006;32:29–33. doi:10.1016/j.ejso.2005.08.010
18. Cutuli B. Assessment of axillary lymph node involvement in small breast cancer: analysis of 893 cases. Clin Breast Cancer. 2001;2:59–65; discussion 66. doi:10.3816/CBC.2001.n.012
19. Andea AA, Bouwman D. Correlation of tumor volume and surface area with lymph node status in patients with multifocal/multicentric breast carcinoma. Cancer. 2004;100:20–27. doi:10.1002/cncr.11880
20. Olivotto IA, Jackson JS, Mates D, Andersen S, Davidson W, Bryce C. Prediction of axillary lymph node involvement of women with invasive breast carcinoma: a multivariate analysis. Cancer. 1998;83:948–955. doi:10.1002/(SICI)1097-0142(19980901)83:5<948::AID-CANC21>3.0.CO;2-U
21. Ravdin PM, De Laurentis M. Prediction of axillary lymph node status in breast cancer patients by use of prognostic indicators. J Natl Cancer Inst. 1994;86:1771–1775. doi:10.1093/jnci/86.23.1771
22. Xie F, Yang H, Wang S, Zhou B, Tong F. A logistic regression model for predicting axillary lymph node metastases in early breast carcinoma patients. Sensors. 2012;12:9936–9950. doi:10.3390/s120709936
23. Viale G, Zurrida S, Maiorano E, Mazzarol G, Pruneri G, Paganelli G. Predicting the status of axillary sentinel lymph nodes in 4351 patients with invasive breast carcinoma treated in a single institution. *Cancer*. 2005;103:492–500. doi:10.1002/cncr.20809

24. Rivadeneira DE, Simmons RM, Christos PJ, Hanna K, Daly JM. Predictive factors associated with axillary lymph node metastases in T1a and T1b breast carcinomas: analysis in more than 900 patients. *J Am Coll Surg*. 2000;191:1–6; discussion 6–8. doi:10.1016/S1072-7515(00)000310-0

25. Toikkanen S. Invasive lobular carcinoma of the breast has better short- and long-term survival than invasive ductal carcinoma. *Br J Cancer*. 1997;76:1234–1240. doi:10.1038/bjc.1997.540

26. Chung MJ, Lee JH, Kim SH, Suh YJ. Simple Prediction Model of Axillary Lymph Node Positivity After Analyzing Molecular and Clinical Factors in Early Breast Cancer. *Medicine*. 2016;95:e3689. doi:10.1097/MD.000000000003689

27. Bozzetti C, Musolino A, Camisa R, et al. Evaluation of HER-2/neu amplification and other biological markers as predictors of response to neoadjuvant anthracycline-based chemotherapy in primary breast cancer: the role of anthracycline dose intensity. *Am J Clin Oncol*. 2006;29:171–177. doi:10.1097/01.coc.0000204405.96572.f9

28. Petit T, Wilt M, Velten M, et al. Comparative value of tumour grade, hormonal receptors, Ki-67, HER-2 and topoisomerase II alpha status as predictive markers in breast cancer patients treated with neoadjuvant anthracycline-based chemotherapy. *Eur J Cancer*. 2004;40:205–211. doi:10.1016/S0959-8049(03)00675-0

29. José B, Michael WK, Jane VF, et al. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. *J Clin Oncol*. 2007;25:3670–3679. doi:10.1200/ JCO.2006.08.8013

30. Chen J, Chen J, Yang B, et al. Predicting sentinel lymph node metastasis in a Chinese breast cancer population: assessment of an existing nomogram and a new predictive nomogram. *Breast Cancer Res Treat*. 2012;135:839–848. doi:10.1007/s10549-012-2219-x