Zhang *et al.*
World Journal of Surgical Oncology (2025) 23:154
https://doi.org/10.1186/s12957-025-03807-0

World Journal of Surgical Oncology

RESEARCH

Open Access

Preoperative lymph node metastasis risk assessment in invasive micropapillary carcinoma of the breast: development of a machine learning-based predictive model with a web-based calculator

Yan Zhang[1], Nan Wang[2], Yuxin Qiu[1], Yingxiao Jiang[2], Peiyan Qin[2], Xiaoxiao Wang[2], Yang Li[2,3], Xiangdi Meng[2,3,4*] and Furong Hao[2,3*]

Abstract

Background Invasive micropapillary carcinoma (IMPC) is a rare subtype of breast cancer characterized by a high risk of lymph node metastasis (LNM). The study aimed to identify predictors of LNM and to develop a machine learning (ML)-based risk prediction model for patients with breast IMPC.

Methods We retrospectively analyzed a cohort of 229 patients diagnosed with breast IMPC between 2019 and 2021. Patients were randomly assigned to training and test sets in a 7:3 ratio. Independent risk factors for LNM were identified using univariable and multivariable logistic regression analyses. Thirteen ML algorithms were trained and compared to determine the optimal model. Model performance was evaluated using the area under the curve (AUC), calibration plots, and decision curve analysis. Internal validation was performed using 100 iterations of tenfold cross-validation.

Results LNM was present in 158 patients (69%). Tumor size, histological grade, progesterone receptor staining intensity, and lymphovascular invasion were identified as independent predictors of LNM (all $p < 0.05$). Among the 13 ML models, logistic regression (LR) demonstrated the best performance, achieving an AUC of 0.88 in the test set. A nomogram based on the LR model was constructed to facilitate clinical application, showing excellent calibration, clinical utility, and a classification accuracy of 76% (95% confidence interval: 70%–82%). The median AUC across cross-validation iterations was 0.83 (interquartile range: 0.76–0.91).

Conclusions This study identified key predictors of LNM in breast IMPC and developed a well-calibrated nomogram to support individualized treatment decision-making.

Keywords Invasive micropapillary carcinoma, Breast cancer, Lymph node metastasis, Machine learning, Prediction model, Nomogram

*Correspondence:
Xiangdi Meng
xiangdi_meng0211@163.com
Furong Hao
hkc515@163.com
Full list of author information is available at the end of the article

Background

Invasive micropapillary carcinoma (IMPC) is a rare but highly aggressive subtype of breast cancer, accounting for approximately 0.9% to 8.4% of all cases [1–5]. Compared with invasive carcinoma of no special type, IMPC exhibits a higher propensity for lymphovascular invasion (LVI) and lymph node metastasis (LNM) [4, 6–8]. As LNM is a critical determinant of treatment strategy and prognosis in breast cancer, its accurate assessment is essential for guiding both surgical and systemic therapy. Given the aggressive nature of IMPC, more extensive axillary surgery—often including complete axillary lymph node dissection—is commonly recommended to improve locoregional control [9, 10]. However, while this approach may reduce recurrence risk, it also carries a higher risk of surgical morbidity, including postoperative lymphedema and potential overtreatment [11]. These challenges underscore the need for more precise preoperative tools to assess LNM risk in IMPC, enabling a more tailored surgical approach that balances oncologic safety with morbidity reduction.

Although several predictive models for LNM in IMPC have been proposed, their performance remains limited due to reliance on public datasets and omission of critical pathologic variables [12, 13]. Previous studies have identified several factors associated with LNM in IMPC, including LVI, high Ki- 67 index, hormone receptor (HR) positivity, and high histologic grade [4, 5, 9, 14]. Incorporating such preoperative pathologic indicators may enhance the accuracy of LNM prediction. In recent years, machine learning (ML) has been increasingly applied to medical diagnostics and outcome prediction, demonstrating superior performance in risk stratification and clinical decision-making support [12, 15, 16]. Applying ML to preoperative LNM prediction in IMPC may improve risk assessment and facilitate individualized surgical planning.

This study aimed to identify key predictors of LNM in IMPC and to develop a ML–based model for preoperative risk assessment. To enhance its clinical utility, we also developed a web-based calculator to support real-time application in clinical settings.

Materials and methods

Data sources, patient selection, and variables

This retrospective study included patients with breast IMPC who underwent surgical treatment at our institution between 2019 and 2021. The study was conducted in accordance with the principles of the Declaration of Helsinki. Owing to its retrospective nature and the anonymization of all data, the requirement for ethical review and individual informed consent was waived by the Ethics Committee of Weifang People's Hospital.

All patients were histopathologically diagnosed with breast IMPC. Based on definitions from previous studies [8, 17, 18], a tumor was classified as IMPC if any proportion of micropapillary components was present, regardless of the percentage within the tumor. The following exclusion criteria were applied: (1) history of other malignancies; (2) absence of either preoperative or postoperative pathology; (3) lack of sentinel lymph node biopsy (SLNB) or axillary lymph node dissection (ALND); (4) bilateral breast cancer, which was excluded to avoid intertumoral heterogeneity that may interfere with pathological assessment and data interpretation; (5) male patients, due to their low incidence and distinct biological characteristics compared to female breast cancer; and (6) incomplete clinical data.

LNM was determined via SLNB or ALND and defined as the presence of micrometastases (0.2–2 mm) or macrometastases (> 2 mm) in any lymph node [19, 20]. The following clinicopathologic variables were collected: age, clinical T stage (cT), clinical M stage (cM), histologic grade (assessed using the Nottingham grading system) [21], estrogen receptor (ER) status, progesterone receptor (PR) status, human epidermal growth factor receptor 2 (Her- 2) status, LVI (assessed via D2 -40 immunohistochemistry and hematoxylin–eosin staining) [22], Ki- 67 index, tumor suppressor protein p53, and cytokeratin 5/6 (CK5/6) expression.

According to the 2020 American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) Clinical Practice Guideline Focused Update, a tumor was considered ER/PR-negative if fewer than 1% of tumor cell nuclei were immunoreactive; all other cases were considered ER/PR-positive [23]. In light of the high ER and PR positivity rate in breast IMPC [4, 5, 9] and to support individualized model development, ER and PR status was further stratified based on average staining intensity. At our institution, the staining intensity levels were categorized as negative (−), low positive (+), positive (+ +), and strong positive (+ ++) in accordance with the CAP-recommended reporting template for biomarker assessment in breast cancer specimens [24–26]. Her- 2 status was evaluated in accordance with the 2018 ASCO/CAP guidelines. Her- 2 immunohistochemistry (IHC) scores of 0 and 1+ are interpreted as negative, while a score of 3+ is considered positive. Cases with an IHC score of 2+ are classified as equivocal and require additional testing using dual-probe fluorescence in situ hybridization (FISH) to determine Her- 2 amplification status [27]. For analysis, patients were categorized into two groups: Her- 2-negative (IHC 0, 1+, or 2+ without amplification) and Her- 2-positive (IHC 3+ or 2+ with amplification). All pathological and immunohistochemical slides were independently evaluated and confirmed

by two experienced pathologists. These assessments were based on a comprehensive evaluation of the entire tumor, including all histological components, rather than being limited to the IMPC areas. Tumor staging was determined according to the 8th edition of the American Joint Committee on Cancer (AJCC) staging system.

Statistical modeling

Predictor selection

Continuous variables were summarized as medians with interquartile ranges (IQRs), and categorical variables were presented as counts and percentages. The Wilcoxon rank-sum test was used to compare continuous variables, while categorical variables were assessed using the chi-square test or Fisher's exact test, as appropriate. Univariable logistic regression was performed to identify potential risk factors for LNM based on clinicopathological features. Variables with a p-value <0.05 in univariable analysis were included in a multivariable logistic regression model to identify independent predictors. A two-sided p-value <0.05 was considered statistically significant.

Model development and assessment

The dataset was randomly divided into training and test sets in a 7:3 ratio, with 70% of cases used for model development and 30% reserved for validation. Based on independent predictors identified through multivariable analysis, thirteen ML algorithms were applied to construct predictive models. These included logistic regression (LR), support vector machines (SVM) with four kernel functions (linear, polynomial, radial basis function, and sigmoid), random forest, naïve Bayes, decision tree, k-nearest neighbors (KNN), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), adaptive boosting (AdaBoost), and bootstrap aggregating (bagging).

Model performance was evaluated by comparing the area under the receiver operating characteristic (ROC) curve (AUC) across models. The model with the highest AUC on the test set was selected as the optimal model. Calibration was assessed using calibration curves, with closer alignment to the 45-degree reference line indicating better agreement between predicted and observed outcomes. Clinical utility was evaluated via decision curve analysis (DCA), which estimates the net benefit across a range of threshold probabilities. Discrimination performance was quantified using AUC, with values approaching 1.0 indicating superior predictive ability.

To minimize overfitting and account for variation due to random partitioning, we conducted 100 iterations of tenfold cross-validation and computed the median AUC with IQR. The optimal classification threshold was determined using the maximum Youden index derived from the ROC curve. Model performance metrics—including accuracy, sensitivity, and specificity—were calculated using confusion matrix analysis. A web-based version of the final model was developed to support clinical implementation and facilitate user accessibility.

All statistical analyses and model development were conducted using R software (version 4.2.1; https://www.R-project.org).

Results

A total of 229 patients with breast IMPC were included in the study, of whom 69% (158/229) had LNM. There were no significant differences in baseline characteristics between the training and test sets (Table 1). Univariable logistic regression identified tumor size, cT stage, PR status, PR staining intensity, LVI, and histologic grade as potential predictors of LNM (all $p< 0.05$; Additional File 1). These variables were subsequently included in the multivariable logistic regression model. As shown in Fig. 1, tumor size, PR staining intensity, LVI, and histologic grade were confirmed as independent predictors of LNM (all $p< 0.05$) and were therefore used to construct the predictive model.

Figure 2 showed the ROC curves and corresponding AUC values for the 13 ML algorithms evaluated. LR model demonstrated the best performance, with an AUC of 0.88 in the test set, and was selected as the final model. A nomogram was constructed to visualize the LR model, assigning a risk score to each predictor and generating a total score corresponding to the predicted probability of LNM (Fig. 3A).

The calibration curve (Fig. 3B) showed strong agreement between predicted and observed probabilities, indicating good model calibration. DCA demonstrated that use of the nomogram provided greater net clinical benefit across a wide range of threshold probabilities (0.1–0.9) compared to treating all or no patients (Fig. 3C). In 1,000 iterations of cross-validation, the model showed consistently robust discriminatory performance, with a median AUC of 0.83 (IQR: 0.76–0.91) (Fig. 3D). A web-based version of the nomogram was developed and is available at https://dynapp.shinyapps.io/IMPC_LNM/, enabling clinicians to rapidly estimate an individual patient's LNM probability by entering relevant clinical parameters (Fig. 4).

Based on the nomogram, a total risk score was calculated for each patient. The optimal cutoff value of 100.4 was determined using the maximum Youden index derived from the ROC curve. Patients with a score >100.4 were classified as high risk, while those with scores ≤ 100.4 were classified as low risk (Fig. 5A). As shown in the confusion matrix (Fig. 5B), the model achieved a

Table 1 Clinicopathologic features of the patients ($n = 229$)

Characteristics	Overall $n = 229$ (%)	Training $n = 160$ (%)	Test $n = 69$ (%)	*P value*
Age (years), Median [IQR]	51 [45, 60]	52.0 [44, 63]	50.0 [45, 59]	0.321
Tumor size (cm), Median [IQR]	2.0 [1.5, 3.0]	2.0 [1.7, 3.0]	2.0 [1.5, 3.0]	0.483
T stage (AJCC 8 th)				0.134
T1	117 (51.1)	86 (53.8)	31 (44.9)	
T2	101 (44.1)	69 (43.1)	32 (46.4)	
T3	11 (4.8)	5 (3.1)	6 (8.7)	
N stage (AJCC 8 th)				0.368
N0	71 (31.0)	53 (33.1)	18 (26.1)	
N +	158 (69.0)	107 (66.9)	51 (73.9)	
M stage (AJCC 8 th)				0.861
M0	226 (98.7)	158 (98.8)	68 (98.6)	
M1	3 (1.3)	2 (1.2)	1 (1.4)	
TNM stage (AJCC 8 th)				0.677
I	45 (19.7)	34 (21.2)	11 (15.9)	
II	80 (34.9)	57 (35.6)	23 (33.3)	
III	101 (44.1)	67 (41.9)	34 (49.3)	
IV	3 (1.3)	2 (1.2)	1 (1.4)	
ER status				0.289
Negative	26 (11.4)	21 (13.1)	5 (7.2)	
Positive	203 (88.6)	139 (86.9)	64 (92.8)	
ER Intensity				0.200
-	26 (11.4)	21 (13.1)	5 (7.2)	
+	14 (6.1)	12 (7.5)	2 (2.9)	
+ +	57 (24.9)	41 (25.6)	16 (23.2)	
+ + +	132 (57.6)	86 (53.8)	46 (66.7)	
PR status				0.511
Negative	30 (13.1)	23 (14.4)	7 (10.1)	
Positive	199 (86.9)	137 (85.6)	62 (89.9)	
PR Intensity				0.995
-	35 (15.3)	25 (15.6)	10 (14.5)	
+	24 (10.5)	17 (10.6)	7 (10.1)	
+ +	62 (27.1)	43 (26.9)	19 (27.5)	
+ + +	108 (47.2)	75 (46.9)	33 (47.8)	
Her- 2 status				0.952
Negative	157 (68.6)	109 (68.1)	48 (69.6)	
Positive	72 (31.4)	51 (31.9)	21 (30.4)	
Ki- 67 (%), median [IQR]	30 [20, 40]	30 [20, 32]	30 [20, 40]	0.821
Subtype				
Luminal A	17 (7.4)	16 (10.0)	1 (1.4)	0.054
Luminal B	194 (84.7)	130 (81.2)	64 (92.8)	
Her- 2 overexpression	13 (5.7)	9 (5.6)	4 (5.8)	
TNBC	5 (2.2)	5 (3.1)	0 (0.0)	
p53 (%), median [IQR]	5.0 [1.0, 15]	5.0 [1.0, 15]	5.0 [2.0, 25]	0.143
CK5/6				0.583
Negative	218 (95.2)	151 (94.4)	67 (97.1)	
Positive	11 (4.8)	9 (5.6)	2 (2.9)	
LVI				0.449
-	26 (11.4)	16 (10.0)	10 (14.5)	

Table 1 (continued)

Characteristics	Overall *n* = 229 (%)	Training *n* = 160 (%)	Test *n* = 69 (%)	*P value*
+	203 (88.6)	144 (90.0)	59 (85.5)	
Histological grade				0.973
I	21 (9.2)	15 (9.4)	6 (8.7)	
II	160 (69.9)	112 (70.0)	48 (69.6)	
III	48 (21.0)	33 (20.6)	15 (21.7)	

Abbreviations: *LNM* Lymph node metastasis, *ER* Estrogen receptor, *PR* Progesterone receptor, *Her- 2* Human epidermal growth factor receptor- 2, *AJCC* American Joint Committee on Cancer, *TNM* Tumor-node-metastasis, *LVI* Lympho-vascular invasion

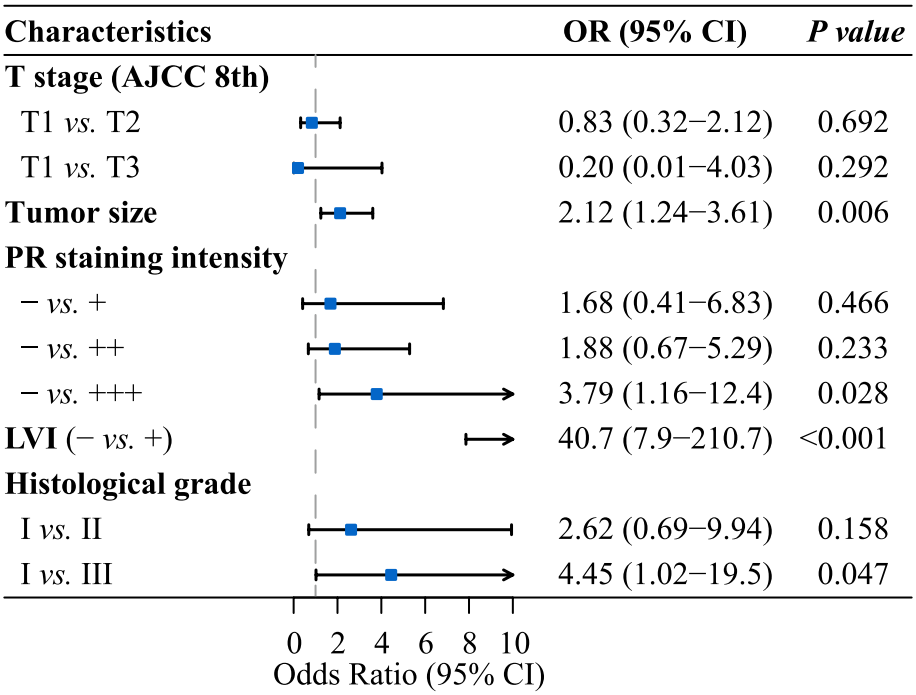

Fig. 1 Multivariable logistic regression forest plots for the candidate predictors. Abbreviations: AJCC, American Joint Committee on Cancer; CI, confidence interval; OR, odds ratios; PR, progesterone receptor; LVI, lympho-vascular invasion

classification accuracy of 76% (95% CI: 70%–82%), with a sensitivity of 62% and specificity of 83%.

Discussion

IMPC was first described by Fisher et al. in 1980 [28] and was formally recognized as a distinct histological subtype of breast cancer by the World Health Organization in 2003 [29]. IMPC often coexists with other histologic subtypes in varying proportions, with pure IMPC accounting for only 0.9% to 2% of all breast cancers [2, 30]. Despite its rarity, numerous studies have highlighted the highly aggressive nature of IMPC [18, 31–33]. Fu Li et al. reported that even when IMPC components comprise less than 10% of the tumor, the malignancy is significantly greater than in tumors without IMPC components [6]. Reported LNM rates in IMPC range from 44 to 85% [34], underscoring its aggressive clinical behavior. In our cohort, the LNM rate was 69%, further confirming the high metastatic potential of IMPC. Accurate preoperative evaluation of regional lymph node involvement is therefore essential to guide appropriate treatment strategies in this patient population. In this study, we identified four independent predictors of LNM—tumor size, PR staining intensity, LVI, and histologic grade. Based on these factors, we developed a nomogram for the preoperative prediction of LNM to support individualized clinical decision-making.

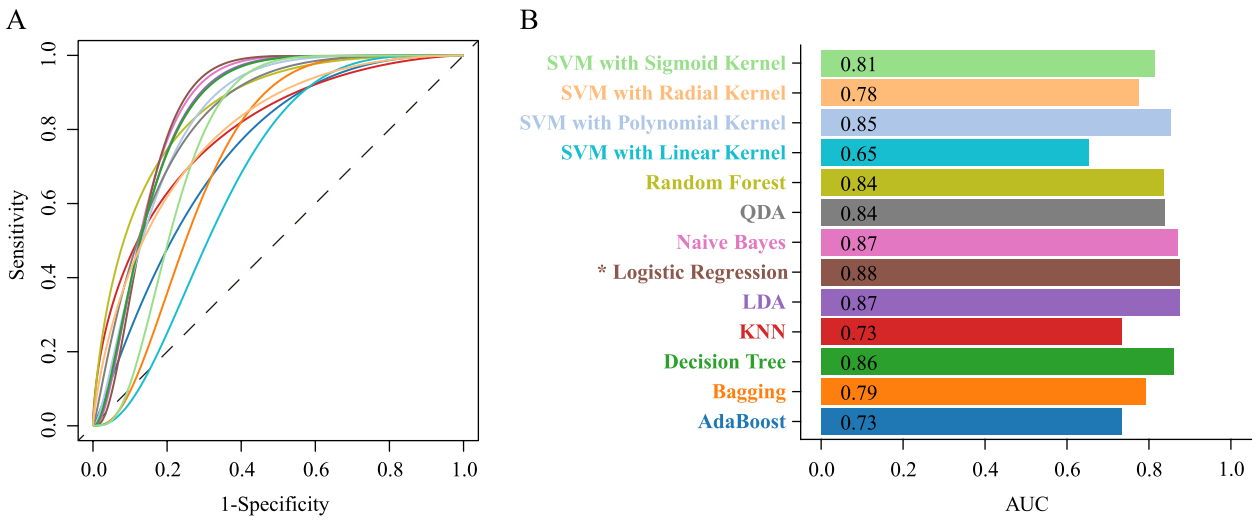

Fig. 2 Comparisons of performance evaluation of predictive models developed via 13 machine learning algorithms. **A** receiver operating characteristic curve of 13 predictive models; **B** The area under the curve values comparisons of the models; Abbreviations: SK, SVM with sigmoid kernel; RK, SVM with radial kernel; PK, SVM with polynomial kernel; LK, SVM with linear kernel; LR, logistic regression; RF, random forest; QDA, quadratic discriminant analysis; LDA, linear discriminant analysis; NB, naive Bayesian; KNN, K-nearest neighbor; DT, decision tree; AdaBoost, adaptive boosting; Bagging, bootstrap aggregating

LVI was identified as an independent predictor and, to our knowledge, was incorporated for the first time into a predictive model for LNM in patients with IMPC. Guo et al. identified lymphatic vessel density and lymphocytic infiltration in IMPC as key factors influencing LNM [7]. Their findings suggested that the presence of lymphatic infiltration increased the probability of LNM in IMPC and that specific chemokines or cytokines might play a regulatory role in this process. In 2009, they demonstrated that adhesion between cancer cells and lymphatic endothelial cells expressing stromal cell–derived factor 1 (SDF- 1) was a critical step in LNM development [35]. Gong et al. reported that loss or reduction of CD44 immunoreactivity was common in IMPC and associated with lymph node positivity [36]. Similarly, the absence of CD44 was more frequently observed in IMPC tumors with LVI and appeared to promote tumor cell infiltration into lymphatic vessels [37]. Collectively, these findings support a strong association between LVI and LNM in breast IMPC.

Tumor size, histologic grade, and HR positivity have frequently been identified as significant risk factors for breast LNM [38, 39]. Guo et al. reported that high histologic grade was associated with the extent of LNM in patients with IMPC [7], while Jiang et al. identified tumor size as the most influential predictor of LNM based on Shapley value analysis [12]. Although IMPC is often characterized by high ER and PR positivity [30, 33, 40], our study found that ER status was not an independent predictor of LNM. Conversely, PR staining intensity was independently associated with LNM, consistent with prior reports. For instance, Giuseppe et al. observed that PR negativity was significantly associated with a lower risk of sentinel LNM [41], while Ravdin et al. reported a positive correlation between PR concentration and the risk of axillary LNM [42]. However, Ye et al. found that ER status was an independent predictor of LNM in IMPC [13], highlighting ongoing controversy regarding the role of hormone receptors in nodal involvement. The expression of ER and PR in breast cancer is believed to reflect the activity of functional estrogen signaling pathways [43]. Gann et al. demonstrated that tumors lacking both ER and PR were significantly less likely to exhibit LNM than tumors expressing both receptors [44]. ER expression is regulated by three genes: ERα, ERβ, and the membrane-bound G protein–coupled receptor 30 (GPR30). PR is a downstream target of ER, primarily regulated by ERα and dependent on estrogen stimulation [45]. Moreover, PR has been shown to modulate ERα activity by influencing chromatin binding and transcriptional regulation [46]. It has been hypothesized that the absence of PR may reflect a dysfunctional or inactive ER signaling pathway, which is also associated with reduced responsiveness to endocrine therapies such as tamoxifen, a selective estrogen receptor modulator (SERM) [47]. In light of these complex interrelationships, our findings suggest that PR status may serve as a more robust predictor of LNM than ER status in patients with IMPC.

In addition, molecular subtype was included as a variable in this study. However, the results showed no

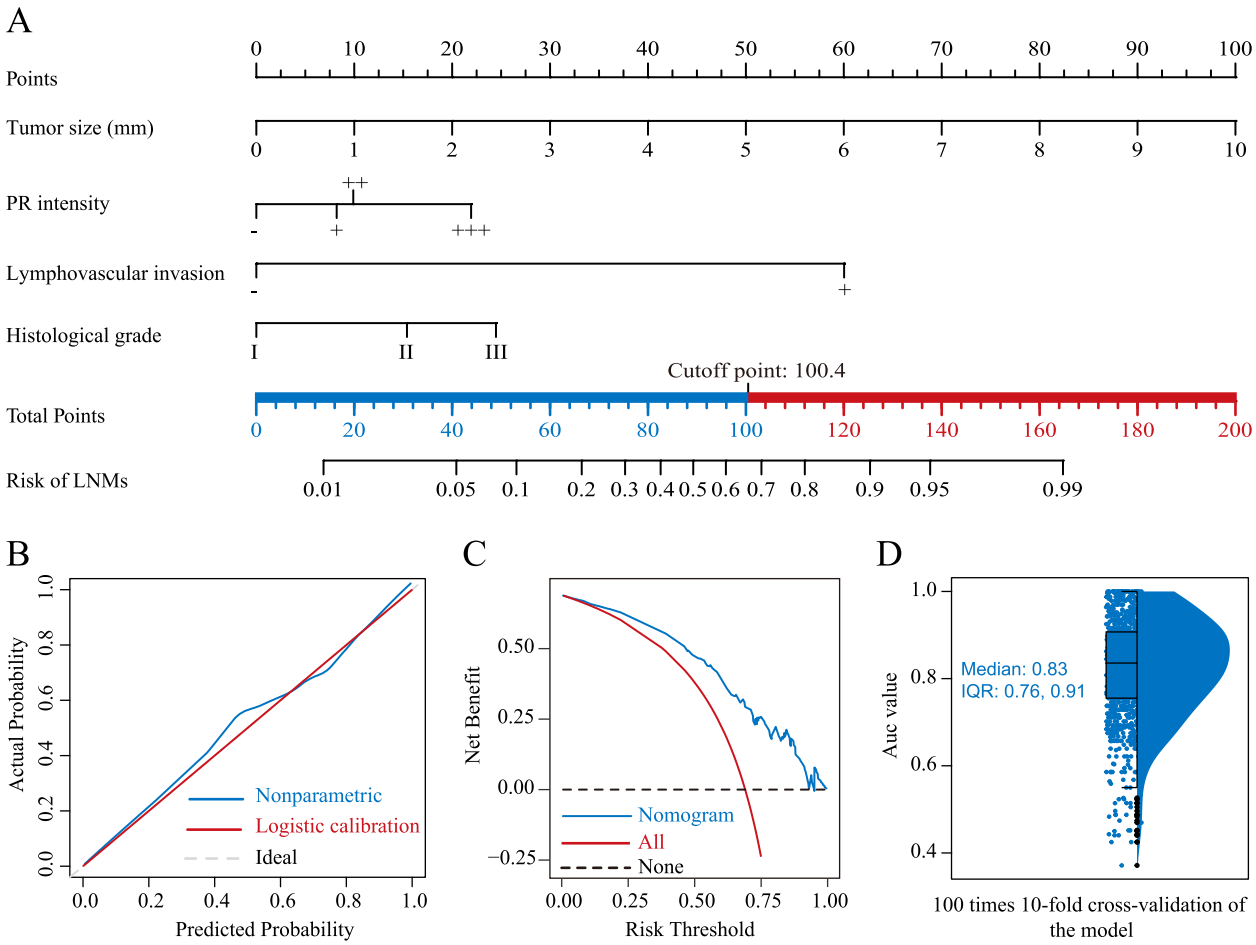

Fig. 3 Model development. **A** nomogram containing independent risk factors for predicting lymph node metastasis (LNM); **B** calibration curve testing model calibration; **C** decision curve analysis assessing the clinical utility; **D** tenfold internal cross-validation for the predictive model; Abbreviations: PR, progesterone receptor; AUC, area under the curve

statistically significant difference in the rate of lymph node metastasis among patients with different molecular subtypes. In contrast, Si et al. reported notable differences in nodal positivity among breast cancer patients with distinct molecular subtypes, with luminal-type tumors demonstrating a stronger association with LNM compared to triple-negative breast cancer (TNBC) [48]. Similarly, Lee et al. found that TNBC subtype was an independent predictor of LNM in breast cancer [49]. Reyal et al. reported that the interaction term between ER and Her- 2 status was an independent predictor of sentinel lymph node positivity, with stronger predictive value than ER status alone [50]. Given the more aggressive behavior of IMPC and the typically high expression of ER, PR, and HER2, it is reasonable to hypothesize that IMPC may exhibit greater tumor heterogeneity. The relationship between molecular subtype and LNM in IMPC remains inconclusive and may be affected by inconsistencies in subtype classification criteria across institutions. Further investigation is warranted to clarify these associations.

The nomogram underwent internal validation and demonstrated favorable predictive performance. The calibration curve indicated good agreement between predicted and observed probabilities. Notably, DCA revealed a wide threshold probability range, suggesting that applying the nomogram to guide clinical decision-making would result in greater net benefit across a broad range of clinical scenarios. The range of AUC values obtained through internal cross-validation further supported the model's stability. Given the importance of identifying patients at varying risk levels, we developed a risk stratification system based on the nomogram. This system demonstrated satisfactory discriminatory ability when compared with nonparametric prediction methods, indicating that the model may offer clinicians a more accurate and individualized reference to inform treatment strategies.

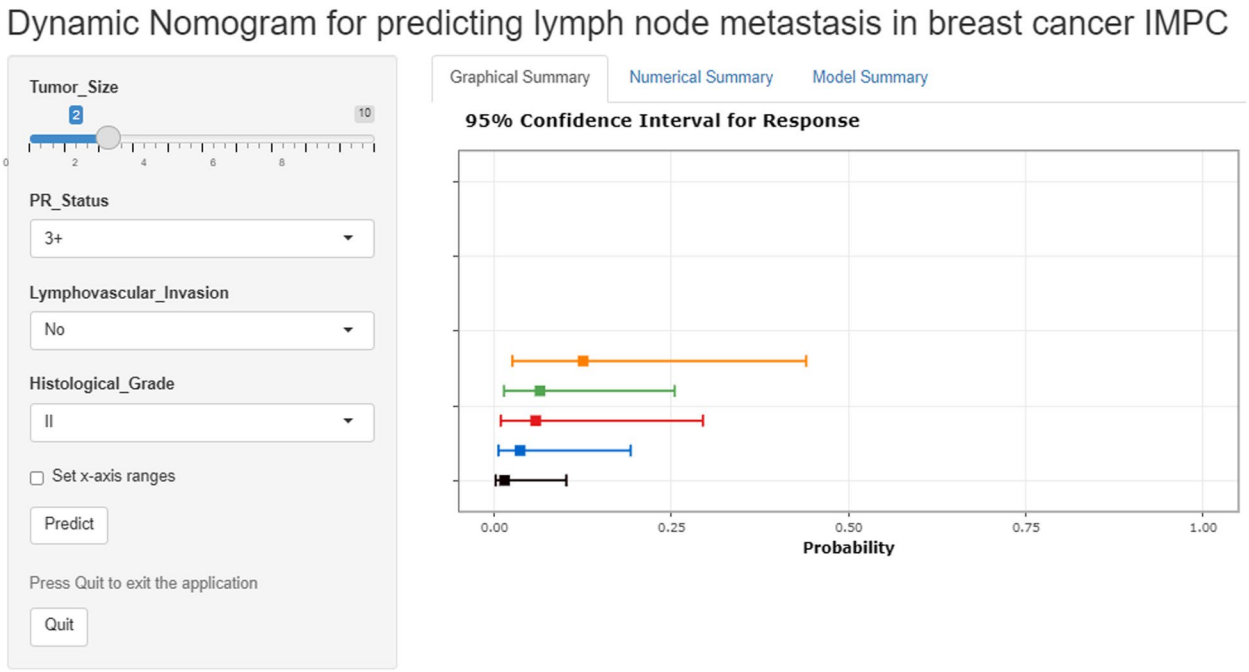

Fig. 4 Web version of the nomogram predicting the probability of lymph node metastasis in IMPC patients (https://dynapp.shinyapps.io/IMPC_LNM/)

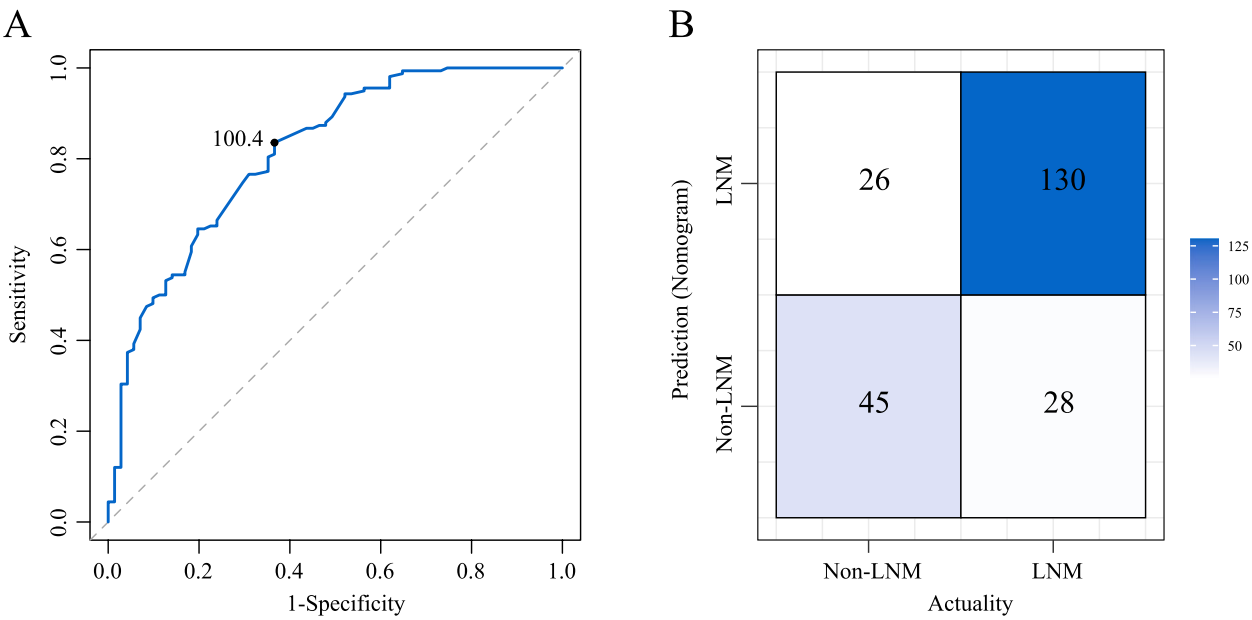

Fig. 5 Application of the model. **A** risk stratification constructed by calculating an optimal threshold from the receiver operating characteristic (ROC) curve; **B** confusion matrix assessing the difference between predicted and actual risk for lymph node metastasis

This study has several limitations. First, as a retrospective analysis, it is subject to inherent selection bias. Second, the model did not incorporate data from other modalities, such as radiomic or genomic features, which may have enhanced its predictive performance. Third, the study did not consider the impact of IMPC proportion on LNM. Previous studies suggest that the aggressive features of IMPC—including LVI, LNM, and high histological grade—are associated with its presence alone. This may explain why most researchers recommend

diagnosing IMPC when it is identified, regardless of its extent. Nonetheless, future studies are needed to explore the specific influence of both IMPC and non-IMPC components on LNM. Fourth, this was a single-center study, and all patients were from a Chinese population, which may limit the generalizability of the findings. Future multicenter, prospective, and multiethnic studies are warranted to further validate and refine the proposed model.

Conclusion

Tumor size, histologic grade, PR staining intensity, and LVI were identified as significant predictors of LNM in patients with breast IMPC. Based on these factors, we developed a logistic regression–based nomogram to estimate the preoperative risk of LNM. With further validation in multicenter, prospective cohorts, this model may serve as a valuable tool to support individualized treatment planning and improve clinical decision-making.

Abbreviations

AdaBoost	Adaptive boosting
AJCC	American Joint Committee on Cancer
ALND	Axillary lymph node dissection
ASCO	American Society of Clinical Oncology
AUC	Area under the curve
CAP	College of American Pathologists
CI	Confidence interval
CK5/6	Cytokeratin 5/6
DCA	Decision curve analysis
ER	Estrogen receptor
Her- 2	Human epidermal growth factor receptor 2
HR	Hormone receptor
IHC	Immunohistochemistry
IMPC	Invasive micropapillary carcinoma
IQR	Interquartile range
KNN	K-nearest neighbors
LDA	Linear discriminant analysis
LNM	Lymph node metastasis
LR	Logistic regression
LVI	Lymphovascular invasion
ML	Machine learning
PR	Progesterone receptor
QDA	Quadratic discriminant analysis
ROC	Receiver operating characteristic
SLNB	Sentinel lymph node biopsy

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12957-025-03807-0.

Supplementary Material 1.

Acknowledgements

None.

Clinical trial number

Not applicable.

Authors' contributions

YZ, XM and FH analyzed and interpreted the IMPC and predictive model. NW, YQ, PQ, YJ and XW collected the study data. YZ, XM, YL and FH participated in the writing of the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the Development Project in Science and Technology of Weifang (Soft Science) (Weifang Science and Technology Bureau, Grant Number: 2022RKX015). The funders had no role in study design, data collection and analysis, interpretation of data and preparation of the manuscript.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All the procedures followed were in accordance with the Helsinki Declaration of the World Medical Association (as revised in 2013). Given the retrospective nature of the study and the anonymization of all data, institutional review board (IRB) approval was waived, and informed consent from patients was not required.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]School of Clinical Medicine, Shandong Second Medical University, Weifang, China. [2]Department of Radiation Oncology, Weifang People's Hospital, Weifang, China. [3]Weifang Key Laboratory of Radiophysics and Oncological Radiobiology, Weifang, China. [4]Department of Radiation Oncology, Gunma University Graduate School of Medicine, Maebashi, Japan.

Received: 21 February 2025 Accepted: 13 April 2025
Published online: 22 April 2025

References

1. Cui ZQ, Feng JH, Zhao YJ. Clinicopathological features of invasive micropapillary carcinoma of the breast. Oncol Lett. 2015;9:1163–6. https://doi.org/10.3892/ol.2014.2806.
2. Verras GI, Tchabashvili L, Mulita F, Grypari IM, Sourouni S, Panagodimou E, et al. Micropapillary breast carcinoma: From molecular pathogenesis to prognosis. Breast Cancer (Dove Med Press). 2022;14:41–61. https://doi.org/10.2147/bctt.S346301.
3. Chen H, Wu K, Wang M, Wang F, Zhang M, Zhang P. Invasive micropapillary carcinoma of the breast has a better long-term survival than invasive ductal carcinoma of the breast in spite of its aggressive clinical presentations: A comparison based on large population database and case-control analysis. Cancer Med. 2017;6:2775–86. https://doi.org/10.1002/cam4.1227.
4. Meng X, Hao F, Wang N, Qin P, Ju Z, Sun D. Log odds of positive lymph nodes (lodds)-based novel nomogram for survival estimation in patients with invasive micropapillary carcinoma of the breast. BMC Med Res Methodol. 2024;24:90. https://doi.org/10.1186/s12874-024-02218-1.
5. Meng X, Ma H, Yin H, Yin H, Yu L, Liu L, et al. Nomogram predicting the risk of locoregional recurrence after mastectomy for invasive micropapillary carcinoma of the breast. Clin Breast Cancer. 2021;21:e368–76. https://doi.org/10.1016/j.clbc.2020.12.003.
6. Fu L, Ikuo M, Fu XY, Liu TH. Shinichi T [relationship between biologic behavior and morphologic features of invasive micropapillary carcinoma of the breast]. Zhonghua Bing Li Xue Za Zhi. 2004;33:21–5.
7. Guo X, Chen L, Lang R, Fan Y, Zhang X, Fu L. Invasive micropapillary carcinoma of the breast: Association of pathologic features with lymph node

metastasis. Am J Clin Pathol. 2006;126:740–6. https://doi.org/10.1309/axyy-4ajt-mnw6-frmw.
8. Li W, Han Y, Wang C, Guo X, Shen B, Liu F, et al. Precise pathologic diagnosis and individualized treatment improve the outcomes of invasive micropapillary carcinoma of the breast: A 12-year prospective clinical study. Mod Pathol. 2018;31:956–64. https://doi.org/10.1038/s41379-018-0024-8.
9. Yu JI, Choi DH, Huh SJ, Cho EY, Kim K, Chie EK, et al. Differences in prognostic factors and failure patterns between invasive micropapillary carcinoma and carcinoma with micropapillary component versus invasive ductal carcinoma of the breast: Retrospective multicenter case-control study (krog 13–06). Clin Breast Cancer. 2015;15:353-61.e1-2. https://doi.org/10.1016/j.clbc.2015.01.008.
10. Yu JI, Choi DH, Park W, Huh SJ, Cho EY, Lim YH, et al. Differences in prognostic factors and patterns of failure between invasive micropapillary carcinoma and invasive ductal carcinoma of the breast: Matched case-control study. Breast. 2010;19:231–7. https://doi.org/10.1016/j.breast.2010.01.020.
11. Deutsch M, Land S, Begovic M, Sharif S. The incidence of arm edema in women with breast cancer randomized on the national surgical adjuvant breast and bowel project study b-04 to radical mastectomy versus total mastectomy and radiotherapy versus total mastectomy alone. Int J Radiat Oncol Biol Phys. 2008;70:1020–4. https://doi.org/10.1016/j.ijrobp.2007.07.2376.
12. Jiang C, Xiu Y, Qiao K, Yu X, Zhang S, Huang Y. Prediction of lymph node metastasis in patients with breast invasive micropapillary carcinoma based on machine learning and shapley additive explanations framework. Front Oncol. 2022;12:981059. https://doi.org/10.3389/fonc.2022.981059.
13. Ye FG, Xia C, Ma D, Lin PY, Hu X, Shao ZM. Nomogram for predicting preoperative lymph node involvement in patients with invasive micropapillary carcinoma of breast: A seer population-based study. BMC Cancer. 2018;18:1085. https://doi.org/10.1186/s12885-018-4982-5.
14. Luna-Moré S, de los Santos F, Bretón JJ, Cañadas MA. Estrogen and progesterone receptors, c-erbb-2, p53, and bcl-2 in thirty-three invasive micropapillary breast carcinomas. Pathol Res Pract. 1996;192:27–32. https://doi.org/10.1016/s0344-0338(96)80126-9.
15. Meng X, Ju Z, Sakai M, Li Y, Musha A, Kubo N, et al. Normal tissue complication probability model for acute oral mucositis in patients with head and neck cancer undergoing carbon ion radiation therapy based on dosimetry, radiomics, and dosiomics. Radiother Oncol. 2025;204:110709. https://doi.org/10.1016/j.radonc.2025.110709.
16. Duan H, Zhang Y, Qiu H, Fu X, Liu C, Zang X, et al. Machine learning-based prediction model for distant metastasis of breast cancer. Comput Biol Med. 2024;169:107943. https://doi.org/10.1016/j.compbiomed.2024.107943.
17. Walsh MM, Bleiweiss IJ. Invasive micropapillary carcinoma of the breast: Eighty cases of an underrecognized entity. Hum Pathol. 2001;32:583–9. https://doi.org/10.1053/hupa.2001.24988.
18. Chen L, Fan Y, Lang RG, Guo XJ, Sun YL, Cui LF, et al. Breast carcinoma with micropapillary features: Clinicopathologic study and long-term follow-up of 100 cases. Int J Surg Pathol. 2008;16:155–63. https://doi.org/10.1177/1066896907307047.
19. Wang Q, Lin Y, Ding C, Guan W, Zhang X, Jia J, et al. Multi-modality radiomics model predicts axillary lymph node metastasis of breast cancer using mri and mammography. Eur Radiol. 2024;34:6121–31. https://doi.org/10.1007/s00330-024-10638-2.
20. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast cancer-major changes in the american joint committee on cancer eighth edition cancer staging manual. CA Cancer J Clin. 2017;67:290-303. https://doi.org/10.3322/caac.21393.
21. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up. Histopathology. 1991;19:403-10. https://doi.org/10.1111/j.1365-2559.1991.tb00229.x.
22. Zhang Y, Wang H, Zhao H, He X, Wang Y, Wang H. Prognostic significance and value of further classification of lymphovascular invasion in invasive breast cancer: A retrospective observational study. Breast Cancer Res Treat. 2024;206:397–410. https://doi.org/10.1007/s10549-024-07318-6.
23. Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, et al. Estrogen and progesterone receptor testing in breast cancer: Asco/cap guideline update. J Clin Oncol. 2020;38:1346–66. https://doi.org/10.1200/jco.19.02309.
24. Keskinkilic M, Semiz HS, Yavuzsen T, Oztop I. Is the percentage of hormone receptor positivity in hr+ her2-metastatic breast cancer patients receiving cdk 4/6 inhibitor with endocrine therapy predictive and prognostic? Front Oncol. 2024;14:1378563. https://doi.org/10.3389/fonc.2024.1378563.
25. Fitzgibbons PL, Dillon DA, Alsabeh R, Berman MA, Hayes DF, Hicks DG, et al. Template for reporting results of biomarker testing of specimens from patients with carcinoma of the breast. Arch Pathol Lab Med. 2014;138:595–601. https://doi.org/10.5858/arpa.2013-0566-CP.
26. Elzohery Y, Radwan AH, Gareer SWY, Mamdouh MM, Moaz I, Khalifa AM, et al. Micropapillary breast carcinoma in comparison with invasive duct carcinoma. Does it have an aggressive clinical presentation and an unfavorable prognosis? BMC Cancer. 2024;24:992. https://doi.org/10.1186/s12885-024-12673-0.
27. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of american pathologists clinical practice guideline focused update. J Clin Oncol. 2018;36:2105–22. https://doi.org/10.1200/jco.2018.77.8738.
28. Fisher ER, Palekar AS, Redmond C, Barton B, Fisher B. Pathologic findings from the national surgical adjuvant breast project (protocol no.4). Vi. Invasive papillary cancer. Am J Clin Pathol. 1980;73:313–22. https://doi.org/10.1093/ajcp/73.3.313.
29. Böcker W. who classification of breast tumors and tumors of the female genital organs: Pathology and genetics. Verh Dtsch Ges Pathol. 2002;86:116–9.
30. Yang YL, Liu BB, Zhang X, Fu L. Invasive micropapillary carcinoma of the breast: An update. Arch Pathol Lab Med. 2016;140:799–805. https://doi.org/10.5858/arpa.2016-0040-RA.
31. De la Cruz C, Moriya T, Endoh M, Watanabe M, Takeyama J, Yang M, et al. Invasive micropapillary carcinoma of the breast: Clinicopathological and immunohistochemical study. Pathol Int. 2004;54:90–6. https://doi.org/10.1111/j.1440-1827.2004.01590.x.
32. Cheng LH, Yu XJ, Zhang H, Zhang HJ, Jia Z, Wang XH. Advances in invasive micropapillary carcinoma of the breast research: A review. Medicine (Baltimore). 2024;103:e36631. https://doi.org/10.1097/md.0000000000036631.
33. Shi WB, Yang LJ, Hu X, Zhou J, Zhang Q, Shao ZM. Clinico-pathological features and prognosis of invasive micropapillary carcinoma compared to invasive ductal carcinoma: A population-based study from china. PLoS One. 2014;9:e101390. https://doi.org/10.1371/journal.pone.0101390.
34. Li Y, Liu J, Xu Z, Shang J, Wu S, Zhang M, et al. Construction and validation of a nomogram for predicting the prognosis of patients with lymph node-positive invasive micropapillary carcinoma of the breast: Based on seer database and external validation cohort. Front Oncol. 2023;13:1231302. https://doi.org/10.3389/fonc.2023.1231302.
35. Liu F, Lang R, Wei J, Fan Y, Cui L, Gu F, et al. Increased expression of sdf-1/cxcr4 is associated with lymph node metastasis of invasive micropapillary carcinoma of the breast. Histopathology. 2009;54:741–50. https://doi.org/10.1111/j.1365-2559.2009.03289.x.
36. Gong Y, Sun X, Huo L, Wiley EL, Rao MS. Expression of cell adhesion molecules, cd44s and e-cadherin, and microvessel density in invasive micropapillary carcinoma of the breast. Histopathology. 2005;46:24–30. https://doi.org/10.1111/j.1365-2559.2004.01981.x.
37. Badyal RK, Bal A, Das A, Singh G. Invasive micropapillary carcinoma of the breast: Immunophenotypic analysis and role of cell adhesion molecules (cd44 and e-cadherin) in nodal metastasis. Appl Immunohistochem Mol Morphol. 2016;24:151–8. https://doi.org/10.1097/pai.0000000000000167.
38. Klar M, Jochmann A, Foeldi M, Stumpf M, Gitsch G, Stickeler E, et al. The mskcc nomogram for prediction the likelihood of non-sentinel node involvement in a german breast cancer population. Breast Cancer Res Treat. 2008;112:523–31. https://doi.org/10.1007/s10549-007-9884-1.
39. Murata T, Watase C, Shiino S, Jimbo K, Iwamoto E, Yoshida M, et al. Development and validation of a preoperative scoring system to distinguish between nonadvanced and advanced axillary lymph node metastasis in patients with early-stage breast cancer. Clin Breast Cancer. 2021;21:e302–11. https://doi.org/10.1016/j.clbc.2020.11.008.
40. Tang SL, Yang JQ, Du ZG, Tan QW, Zhou YT, Zhang D, et al. Clinico-pathologic study of invasive micropapillary carcinoma of the breast.

Oncotarget. 2017;8:42455–65. https://doi.org/10.18632/oncotarget.16405.
41. Viale G, Zurrida S, Maiorano E, Mazzarol G, Pruneri G, Paganelli G, et al. 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. https://doi.org/10.1002/cncr.20809.
42. Ravdin PM, De Laurentiis M, Vendely T, Clark GM. Prediction of axillary lymph node status in breast cancer patients by use of prognostic indicators. J Natl Cancer Inst. 1994;86:1771–5. https://doi.org/10.1093/jnci/86.23.1771.
43. Clusan L, Ferrière F, Flouriot G, Pakdel F. A basic review on estrogen receptor signaling pathways in breast cancer. Int J Mol Sci. 2023;24. https://doi.org/10.3390/ijms24076834.
44. Gann PH, Colilla SA, Gapstur SM, Winchester DJ, Winchester DP. Factors associated with axillary lymph node metastasis from breast carcinoma: Descriptive and predictive analyses. Cancer. 1999;86:1511–9. https://doi.org/10.1002/(sici)1097-0142(19991015)86:8%3c1511::aid-cncr18%3e3.0.co;2-d.
45. Li Z, Wei H, Li S, Wu P, Mao X. The role of progesterone receptors in breast cancer. Drug Des Devel Ther. 2022;16:305–14. https://doi.org/10.2147/dddt.S336643.
46. Mohammed H, Russell IA, Stark R, Rueda OM, Hickey TE, Tarulli GA, et al. Progesterone receptor modulates erα action in breast cancer. Nature. 2015;523:313–7. https://doi.org/10.1038/nature14583.
47. Cui X, Schiff R, Arpino G, Osborne CK, Lee AV. Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy. J Clin Oncol. 2005;23:7721–35. https://doi.org/10.1200/jco.2005.09.004.
48. Si C, Jin Y, Wang H, Zou Q. Association between molecular subtypes and lymph node status in invasive breast cancer. Int J Clin Exp Pathol. 2014;7:6800–6.
49. Lee JH, Kim SH, Suh YJ, Shim BY, Kim HK. Predictors of axillary lymph node metastases (alnm) in a korean population with t1–2 breast carcinoma: Triple negative breast cancer has a high incidence of alnm irrespective of the tumor size. Cancer Res Treat. 2010;42:30–6. https://doi.org/10.4143/crt.2010.42.1.30.
50. Reyal F, Rouzier R, Depont-Hazelzet B, Bollet MA, Pierga JY, Alran S, et al. The molecular subtype classification is a determinant of sentinel node positivity in early breast carcinoma. PLoS One. 2011;6:e20297. https://doi.org/10.1371/journal.pone.0020297.

Publisher's Note