Diagnostic Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

The development and validation of a risk stratification system for assessing axillary status after neoadjuvant therapy in node-positive breast cancer: a multicenter, prospective, observational study

Jia-Xin Huang, MD[a,b], Jing-Si Mei, MD[c], Fei Chen, MD[d], Jia-Hui Huang, PhD[e], Yu-Ting Tan, MD[c], Yi-Wen Wu, MsD[d], Feng-Tao Liu, MD[c], Shao-Dong Qiu, MD[d], Cai-Gou Shi, MD[f], Yao Lu, PhD[g], Xue-Yan Wang, MD[b], Gui-Ling Huang, MD[b], Yu-Ting Zhang, MD[b], Min-Shan Chen, MD, PhD[a], Xiao-Qinsg Pei, MD, PhD[b,*]

Objective: It is not clear which procedure is most optimal for axilla after neoadjuvant therapy (NAT) in node-positive breast cancer patients. Accurately identifying patients with axillary pathologic complete response (pCR) is crucial to minimize the overtreatment of axilla. This study was designed to develop a risk stratification model for axillary pCR.

Methods: In this multicenter, prospective, observational study, node-positive breast cancer patients who received NAT followed by axillary lymph node dissection (ALND) were enrolled between June 2021 and April 2024. We assessed the performance of breast shear wave elastography (SWE) utilizing virtual touch imaging quantification in determining axillary status across ultrasound (US) nodal stages following NAT. A predictive model incorporating axilla US nodal stage and breast SWE was developed using multivariate logistic regression analysis. Last, a simplified risk score was developed based on the calculated prediction probability from this model and validated in the external test cohort.

Results: The axillary pCR rates were 52.53% in the training cohort (n = 257) and 51.79% in the external test cohorts (n = 195). Approximately 21.67% of US N0 cases were false negatives; 42.35% of US N1 cases were false positives. With SWE, the false negative rate was 11.53% in US N0 patients and false positive rate was 22.22% in US N1 patients. The model based on dual-modality US demonstrated strong discriminatory ability (AUC, 0.93), precise calibration (slope of calibration curve, 0.99), and practical clinical utility (probability threshold, 4.5–94.5%); the percentages of accuracy, sensitivity, and specificity were 87.94%, 88.52%, and 87.41%, respectively. Patients scoring 1 demonstrated a low axillary non-pCR rate (5.21%–6.97%), potentially reducing unnecessary ALND rate (17.12%–24.10%).

Conclusions: The risk stratification model integrating axilla US and breast SWE demonstrated good performance for assessing axillary status after NAT in node-positive breast cancer and might provide guidance for less aggressive management for specific individuals.

Keywords: breast neoplasm, elasticity imaging techniques, lymph node, neoadjuvant therapy

Introduction

For breast cancer patients with lymph node (LN) metastasis, neoadjuvant therapy (NAT) can reduce the tumor burden in both breast tumors and axillary LNs[1]. Approximately half of these patients achieve a pathologic complete response (pCR) in the axillary LNs and experience improved prognosis after NAT[2]. For patients who achieve axillary pCR following

[a]Department of Liver Surgery, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China, [b]Department of Medical Ultrasound, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China, [c]Breast Tumor Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China, [d]Department of Medical Ultrasound, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, [e]Institute of Artificial Intelligence, Guangzhou University, Guangzhou, China, [f]Department of Medical Ultrasound, Liuzhou People's Hospital, Liuzhou, China and [g]School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

Jia-Xin Huang, Jing-Si Mei, Fei Chen, and Jia-Hui Huang contributed equally to this work.

*Corresponding Author. Address: Department of Medical Ultrasound, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou 510060, China. Tel.: +86 13560172209. E-mail: peixq@sysucc.org.cn (X. Pei).

International Journal of Surgery (2025) 111:3731–3741

Received 1 January 2025; Accepted 26 March 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 12 May 2025

http://dx.doi.org/10.1097/JS9.0000000000002391

NAT, the omission of axillary lymph node dissection (ALND) can be considered. Less radical modalities can serve as alternative surgical methods to restage axillary LNs following NAT, which can decrease ALND-associated morbidity and complications[3]. However, the overall false negative rate of sentinel LN biopsy (SLNB) after NAT in node-positive cases is 12%–15%[2,4,5]. In addition, targeted axillary dissection (TAD) is increasingly being considered, but challenges still persist in terms of the identification rate and false negative rate[5-8]. Furthermore, the criteria to select candidates for less radical management as an alternative to ALND after NAT have not been firmly established. Consequently, there has been considerable heterogeneity in surgical approaches to the axilla following NAT in breast cancer patients with confirmed positive LNs[4-10].

Accurate preoperative assessment of nodal response to NAT is crucial in guiding the management of the axilla. According to the American College of Radiology Appropriateness Criteria guidelines, ultrasound (US) is the optimal method to evaluate axillary status following NAT[11]. However, in breast cancer patients with confirmed positive LN, axilla US has limitations in accurately determining residual axillary LN metastasis following NAT[12]. Ultrasonic elastography, a US technology that assesses biomechanical characteristics, holds promise in diagnosing axillary LNs before treatment and assessing breast tumor response to NAT[12-16]. In a meta-analysis of studies on eight imaging modalities for the preoperative detection of axillary LN metastasis in breast cancer patients, elastography exhibited the best performance[17]. However, only a limited number of studies have confirmed the effectiveness of elastic imaging in assessing axillary LNs after NAT[18,19].

Shear wave elastography (SWE) is a recently developed technique that applies acoustic radiation force impulse (ARFI) to generate shear waves. This technique is a highly reproducible method for evaluating the elasticity of breast masses[20]. Virtual touch imaging quantification (VTIQ) is a two-dimensional SWE based on ARFI technology. It not only provides quantitative and qualitative characterization of tissue elasticity but also includes a quality map to assess the reliability of SWE data. SWE based on VITQ has been proven effective in characterizing tumor cellularity and extracellular matrix properties in both breast tumors and axillary LNs following NAT, thereby offering promise in assessing residual disease in these areas after NAT[18]. Nevertheless, the overall false negative of SWE interpretations was 12%–31%, which is unacceptably high for clinical application[18].

Given the varying significance of SLNB in breast cancer patients with different clinical nodal stages, and the current recommendation of axilla conservation after NAT, specifically for cN1 cases[2], we hypothesized that SWE may also exhibit diverse performance across different nodal stages. Thus, the objective of this study was to investigate the value of SWE in determining axillary status in breast cancer patients with different nodal stages observed on US after NAT. We further developed a US risk stratification model using axilla US and breast SWE to predict the probability of residual axillary LN metastasis.

HIGHLIGHTS

- Shear wave elastography (SWE) using virtual touch imaging quantification could improve the diagnosis of axillary lymph node (LN) in breast cancer patients with ultrasound (US) nodal stage of N0-1 after neoadjuvant therapy (NAT).
- The predictive model based on the odds of SWE and US nodal stage demonstrated good discrimination ability, precise calibration, and practical clinical utility for assessing axillary status after NAT in patients with node-positive breast cancer.
- Patients with a risk score of 1 (US N0 with SWV_{mean} <2.77 m/s or US N1 with SWV_{mean} <1.92 m/s) were least likely to have residual axillary metastasis and may be candidates for less radical management, which could potentially reduce the rate of unnecessary axillary LN dissection and increase the rate of final benefit.

Methods

Study design

This multicenter, prospective, observational study was approved by the ethics committee of the institutional review board (number: B2022-373-X01) and registered in the Chinese Clinical Trial Registry http://www.chictr.org/cn/, number: ChiCTR2400085035). Written informed consent for study participation was obtained from all patients. The study has been reported in accordance with the STARD (Standards for the Reporting of Diagnostic Accuracy Studies) criteria[21].

The overall design of this study is shown in Fig. 1. The pathological outcomes of axillary LNs stratified across US nodal stage following NAT was examined. Then, the efficacy of breast SWE in predicting axillary status across various US nodal stages was evaluated. Following this, a predictive model incorporating axilla US nodal stage and breast SWE was formulated and evaluated. Finally, based on the probability of residual axillary LN metastasis, a risk stratification system was developed and externally validated.

Study population

A flowchart of the selection of the study population is shown in Figure 2. We consecutively recruited breast cancer patients with confirmed axillary LN metastasis undergoing NAT from center A (Sun Yat-Sen University Cancer Center). The patient inclusion and exclusion workflow are shown in Supplementary eMethod 1 http://links.lww.com/JS9/E116. Finally, a total of 257 patients from center A were included as the training cohort between June 2021 and July 2023, which was utilized for data analysis and model development.

Between May 2023 and April 2024, patients were prospectively recruited from center B (Sun Yat-Sen Memorial Hospital) and center C (The Second Affiliated Hospital of Guangzhou Medical University). After applying the same inclusion and exclusion criteria, 195 patients were identified as the external test cohort for prospectively testing the risk stratification model to assess its performance and generalizability to other populations.

Image acquisition and interpretation

After the completion of NAT and 1 or 2 days before surgery, patients underwent US examinations, including conventional US and SWE. During the US examination, suspicious US features of

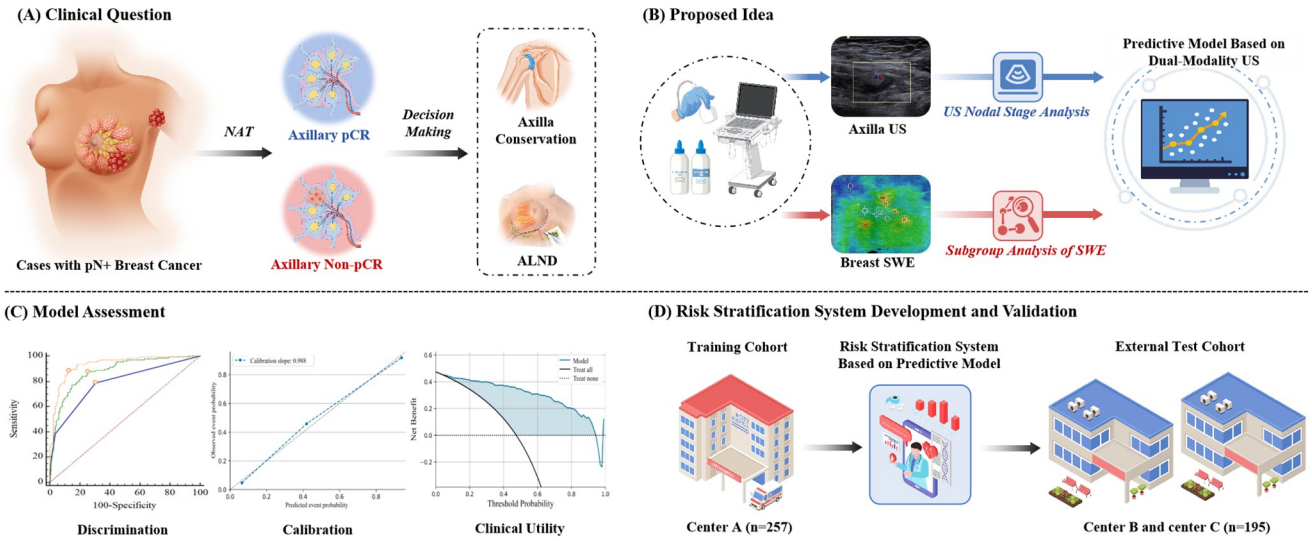

Figure 1. Overview of the study design. (A) Accurately predicting axillary pCR in breast cancer patients before surgery can help clinicians minimize overtreatment of the axilla after NAT. (B) Our hypothesis was that SWE may show varying performance across different nodal stages, and we aimed to create a predictive model using axilla US and breast SWE. (C) We evaluated the discrimination, calibration, and clinical utility of this predictive model. (D) Based on the probability of residual axillary lymph node metastasis following NAT calculated from this predictive model, a risk stratification system was developed and externally tested. Abbreviations: pN+, pathologically positive node; NAT, neoadjuvant therapy; pCR, pathological complete response; ALND, axillary lymph node dissection; US, ultrasound; SWE, shear wave elastography.

axillary LNs and the number of positive LNs on US were recorded[22]. The nodal stage on axilla US was defined with reference to the AJCC pathological nodal staging criteria[23]. Following the conventional US examination, SWE based on VTIQ was performed three times at the plane of the breast tumor's maximal diameter. First, a quality map was first obtained to assess the reliability of the shear wave data. Second, a velocity map was obtained to measure shear wave velocity (SWV) values. Average SWV value was obtained for subsequent analysis[18,24,25]. The detailed procedures for image acquisition and interpretation are

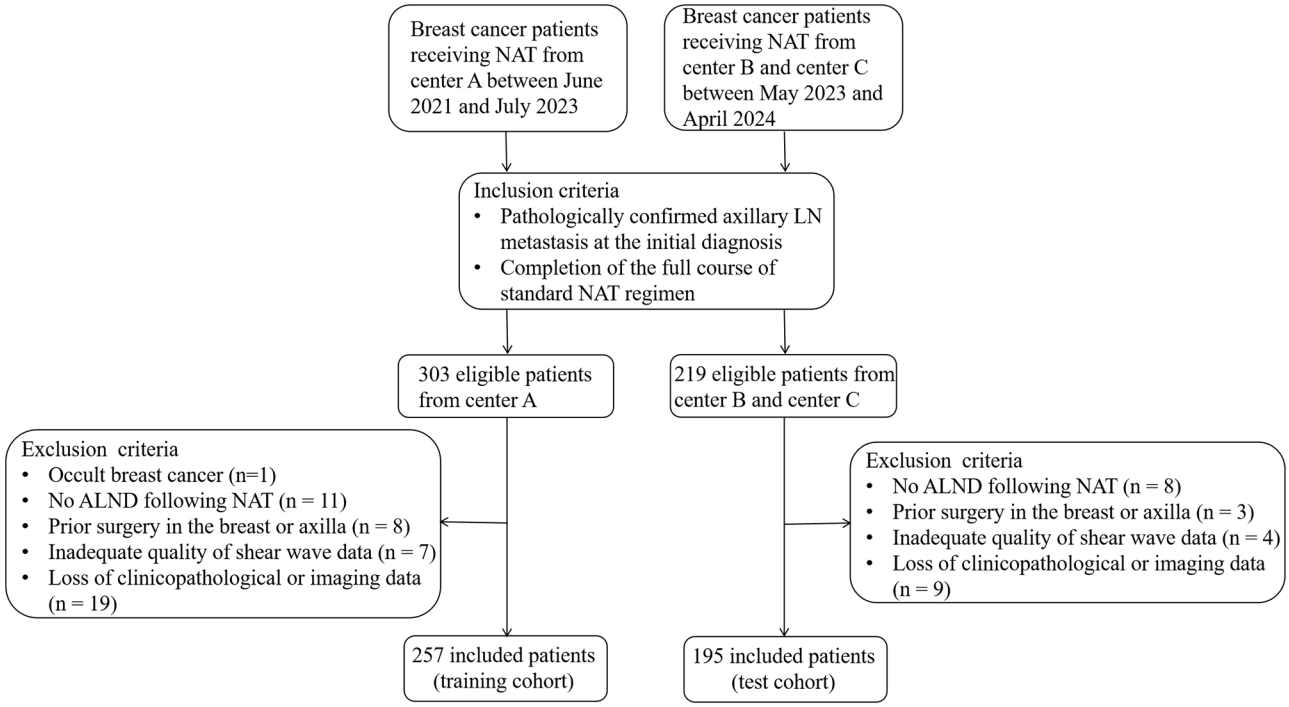

Figure 2. Flowchart of the study population. Abbreviations: NAT, neoadjuvant therapy; LN, lymph node; ALND, axillary lymph node dissection.

provided in Supplementary eMethod 2 http://links.lww.com/JS9/E116.

Pathological evaluation

Prior to NAT, breast cancer and axillary LN metastasis were confirmed through pathological examination of samples obtained via core needle biopsy. Hormone receptor (HR) including estrogen receptor (ER) status and progesterone receptor (PR) status, human epidermal growth factor receptor 2 (HER2) status, and Ki-67 proliferation index were determined by immunohistochemistry and fluorescence in situ hybridization[26]. Based on the above receptor status, all the patients were categorized into three molecular subtypes as follows: (i) HR+HER2−; (ii) HR±HER2+; (iii) HR−HER2−[27].

Following the completion of NAT, all included patients received ALND in conjunction with either mastectomy or breast-conserving surgery. During ALND, LNs in axillary levels I/II were routinely removed; for patients with extensive axillary LNs involvement or evident invasion of axillary level II/III LNs, dissection of axillary level III LNs was performed. All axillary LN specimens were stained with hematoxylin and eosin for the detection of malignant cells and identification of pathologically positive LNs. The location and number of positive LNs were recorded in the pathological report. An axillary pCR was defined as the absence of malignant cells in all surgically resected axillary LNs from a patient. The pathological results of ALND were considered as the gold standard for evaluating the accuracy and effectiveness of the model developed in this study.

Data collection

Clinicopathological data, including age, menopausal status, clinic tumor and nodal stage before treatment, ER status, PR status, HER2 status, Ki-67 proliferation index, and surgical-pathologic results, were obtained from medical records. The nodal stage following NAT observed on axilla US was collected from imaging records. The average SWV value was calculated from breast SWE images[18,24,25].

Statistical analysis

Continuous quantitative variables were compared using either a t-test or Mann–Whitney U test. Categorical variables were analyzed using either the $\chi 2$ test or Fisher exact test. The Spearman's rank test was used to evaluate the intra-observer agreement and inter-observer agreement of both axilla US nodal stage and breast SWE. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the discriminative ability of breast SWE in predicting axillary status after NAT; the specific SWE performances for different US nodal stages were assessed. A multivariate logistic regression (LR) analysis with forward selection (entry P-value = 0.05; removal P-value = 0.10) was performed to identify independent risk factors associated with axillary response to NAT. These factors were integrated to develop a predictive model for residual axillary LNs metastasis. The model incorporated axilla US nodal stage and breast SWE measurement. The formula for this combined predictive model was

$$\log\left(\frac{P}{1-P}\right) = -6.749 + 1.650 * \mathrm{N1(Yes = 1; No = 0)} + 3.539 * \mathrm{N2(Yes = 1; No = 0)} + 4.296 * \mathrm{N3(Yes = 1; No = 0)} + 1.931 * \mathrm{SWV_{mean}}$$

P represents the estimated probability of residual metastasis in axillary LNs after NAT.

The discriminatory ability of the predictive model was assessed by ROC curve analysis. The DeLong test was used for pairwise comparison of the area under the ROC curve (AUC). A calibration curve was developed to show the association between the predictions and observations. The clinical practicability of the model was evaluated using decision curve analysis. Based on the LR formula derived from this predictive model, the risk probability was calculated and categorized into five simplified risk scores[28]. A risk stratification system was subsequently developed to estimate the axillary non-pCR rate for each risk score, enabling a more precise assessment of residual disease likelihood. Sample size was estimated using PASS version 2024 (Supplementary eTable 1 http://links.lww.com/JS9/E116). The other statistical analyses were performed using SPSS version 25.0, MedCalc version 19.8, and Python version 3.12.1.

Results

Clinicopathologic characteristics

The clinicopathological characteristics of the study population are listed in Table 1. A total of 452 eligible patients were included in this study. The average patient age was 47.19 ± 10.51 years (range: 25–74 years). All included patients received ALND, with the number of cleared axillary LNs ranging from 9 to 38 (mean ± SD, 17.86 ± 7.54). Among them, 257 were assigned to the training cohort, with a mean age of 46.39 ± 10.26 years (range: 25–70 years). In the training cohort, 135 (52.53%) achieved axillary pCR and 122 (47.47%) had residual metastasis in axillary LNs following NAT. The external test cohort consisted of 195 patients, with a mean age of 48.25 ± 10.78 years (range: 27–74 years). In the test cohort, a total of 101 (51.79%) achieved axillary pCR and 94 (48.21%) had residual metastasis. There was no significant difference in age, breast cancer histologic type, ER expression, HER2 expression, molecular subtype or axillary response between the training cohort and the external test cohort. However, significant difference was observed in clinical stage pre-NAT, PR, and Ki-67 expression between the two cohorts ($P < 0.05$).

As shown in Supplementary eTable 2 http://links.lww.com/JS9/E116, in the training cohort, compared with patients with axillary pCR, patients with residual nodal metastasis had a significantly higher clinical nodal stage at the initial diagnosis, positive expression of ER, negative expression of HER2, and low expression of Ki-67. Similar to results in the training cohort, patients with residual nodal metastasis in the test cohort had significantly higher clinical nodal stage, positive ER, negative HER2, and low Ki-67 expression.

Agreement for axilla US and breast SWE

The intraclass correlation coefficient (ICC) values for axilla US nodal stage were 0.91 (95% confidence interval (CI), 0.77–0.94) for intra-observer agreement and 0.84 (95% CI, 0.72–0.90) for inter-observer agreement ($P < 0.001$), indicating high levels of both intra- and inter-observer agreement. The intra-observer agreement for breast $\mathrm{SWV_{mean}}$ value (ICC, 0.90; 95% CI, 0.75–0.93) was excellent, and the inter-observer agreement for breast $\mathrm{SWV_{mean}}$ value (ICC, 0.82; 95% CI, 0.73–0.88) was good ($P < 0.001$).

Table 1

Clinicopathological characteristics of the study population

Characteristics	Total	Training cohort (*n* = 257)	External test cohort (*n* = 195)	*P* value
Age, years (mean ± SD)	47.19 ± 10.51	46.39 ± 10.26	48.25 ± 10.78	0.063
Histologic type, *n* (%)				0.299
Invasive ductal carcinoma	420	236 (56.19)	184 (43.81)	
Others	32	21 (65.63)	11 (34.38)	
cT stage, *n* (%)				0.005
1	45	33 (73.33)	12 (26.67)	
2	240	135 (56.25)	105 (43.75)	
3	95	59 (62.11)	36 (37.89)	
4	72	30 (41.67)	42 (58.33)	
cN stage, *n* (%)				<0.001
1	239	98 (41.00)	141 (59.00)	
2	132	83 (62.88)	49 (37.12)	
3	81	76 (93.83)	5 (6.17)	
ER, *n* (%)				0.820
Negative	137	79 (57.66)	58 (42.34)	
Positive	315	178 (56.51)	137 (43.49)	
PR, *n* (%)				0.015
Negative	182	116 (63.74)	66 (36.26)	
Positive	270	141 (52.22)	129 (47.78)	
HER2, *n* (%)				0.262
Negative	239	130 (54.39)	109 (45.61)	
Positive	213	127 (59.62)	86 (40.38)	
Ki-67, *n* (%)				0.011
≤14%	55	40 (72.73)	15 (27.27)	
>14%	397	217 (54.66)	180 (45.34)	
Molecular subtype, *n* (%)				0.439
HR+HER2−	188	103 (54.79)	85 (45.21)	
HR±HER2+	212	127 (59.91)	85 (40.09)	
HR−HER2−	52	27 (51.92)	25 (48.08)	
Axillary response, *n* (%)				0.877
Axillary pCR	236	135 (57.20)	101 (42.80)	
Axillary non-pCR	216	122 (56.48)	94 (43.52)	

Data are shown as *n* (%) unless otherwise indicated.

Abbreviations: cT, clinic tumor; cN, clinic node; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; pCR, pathologic complete response.

Subgroup analysis stratified across US nodal stages

The overall accuracy of axilla US in diagnosing axillary status after NAT was 73.93% (95% CI, 65.37%–81.32%), with a sensitivity of 78.69% (95% CI, 70.43%–85.62%) and a specificity of 69.63% (95% CI, 61.13%–77.24%). The surgical-pathological results for different US nodal stages following NAT are presented in Table 2. The axillary pCR rate for cases with US stage of N0 was 78.33% (94/120), and false negative results were identified in 21.67% (26/120) of patients. Among patients staged as N1 with US, 57.65% (49/85) had pathology-confirmed residual axillary LNs metastasis, and 42.35% (36/85) had false positive results. For patients with US stages of N2 and N3, the axillary pCR rates (false positive results) were 12.90% (4/31) and 4.76% (1/21), respectively.

Table 2

Surgical-pathological results for different US nodal stages after NAT (*n* = 257)

US nodal stage	No. of patients	Axillary pCR, *n* (%)	Axillary non-pCR, *n* (%)	No. of false negative cases	No. of false positive cases
N0	120	94 (78.33)	26 (21.67)	26	0
N1	85	36 (42.35)	49 (57.65)	0	36
N2	31	4 (12.90)	27 (87.10)	0	4
N3	21	1 (4.76)	20 (95.24)	0	1

Abbreviations: US, ultrasound; NAT, neoadjuvant therapy; pCR, pathologic complete response.

SWE performance stratified across US nodal stages

The quantitative VTIQ parameter used to assess residual axillary LN disease is the SWV_{mean} value of breast lesion in our study. The AUC of breast SWE in determining axillary status after NAT was 0.88 (95% CI, 0.84–0.92), yielding an overall accuracy of 80.94% (95% CI, 77.43%–87.16%), a sensitivity of 87.70% (95% CI, 80.52%–93.04%), and a specificity of 74.81% (95% CI, 66.61%–81.90%). As presented in Table 3, in patients with US N0 after NAT, the use of breast SWE reduced the false negative rate to 11.54% (3/26). For patients with US N1 after NAT, the false positive rate decreased to 22.22% (8/36) with breast SWE. For patients with US N2 following NAT, the reduction in false positive results was offset by an increase in false negative results (*n* = 4). For patients with US N3 following NAT, breast SWE resulted in false positive results (*n* = 1) and two cases with false negative results.

Performance of predictive model based on dual-modality US

The odds of residual metastasis in axillary LNs following NAT were significantly improved for cases with higher US

Table 3

SWE performance stratified across US nodal stages (*n* = 257)

US nodal stage	No. of patients	Accuracy (%)	Sensitivity (%)	Specificity (%)	No. of false negative cases	No. of false positive cases
N0	120	76.67 (92/120)	88.46 (23/26)	73.40 (69/94)	3	25
N1	85	83.53 (71/85)	87.76 (43/49)	77.78 (28/36)	6	8
N2	31	87.10 (27/31)	85.19 (23/27)	100.00 (4/4)	4	0
N3	21	85.71 (18/21)	90.00 (18/20)	0.00 (0/1)	2	1

Abbreviations: SWE, shear wave elastography; US, ultrasound.

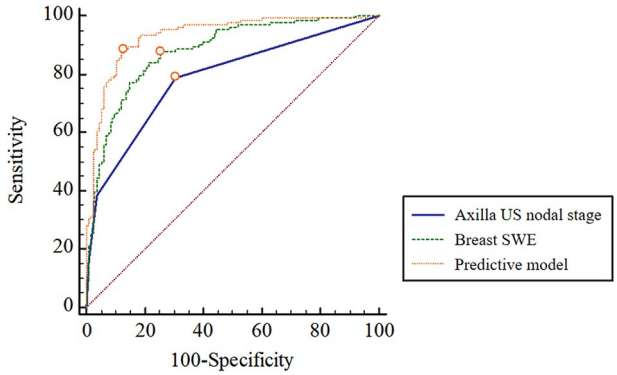

Figure 3. ROC curves of axilla US nodal stage, breast SWE, and the predictive model based on dual-modality US for predicting axillary residual disease following NAT in breast cancer patients. Abbreviations: ROC, receiver operating characteristic; US, ultrasound; SWE, shear wave elastography; NAT, neoadjuvant therapy.

nodal stage, with odds ratio (OR) of 5.21 (95% CI, 2.28–11.93), 34.42 (95% CI, 8.86–70.73) and 73.40 (95% CI, 8.10–155.46) for N1, N2 and N3, respectively, and a higher breast SWV_{mean} with OR of 6.90 (95% CI, 3.95–12.04). The predictive model was developed using multivariate LR analysis, incorporating US nodal stage and breast SWV_{mean} as key variables, and it showed good discrimination ability (AUC, 0.93; 95% CI, 0.90–0.96), with an overall accuracy of 87.94% (95% CI, 80.93%–93.00%), sensitivity of 88.52% (95% CI, 81.5%–93.6%), and specificity of 87.41% (95% CI, 80.6%–92.5%). This model significantly outperformed axilla US or breast SWE alone ($P < 0.01$, Delong's test), as depicted in Figure 3. The *P*-value obtained using the Hosmer–Lemeshow goodness-of-fit test was 0.062, indicating a good fit of the model. Calibration curve analysis showed good agreement between the observations and predictions for axillary LN status after NAT (Fig. 4A). Decision curve analysis showed that clinical decision making according to the risk model offered superior overall benefit to the all-or-none strategy when the probability threshold was between 4.5% and 94.5% (Fig. 4B).

The performance of the predictive model integrating axilla US nodal stage and breast SWE stratified by US nodal stages is shown in Table 4. Compared with axilla US diagnosis, the predictive model demonstrated markedly improved accuracy for cases with US N0 (from 78.33% to 88.33%) and cases with US N1 (from 57.65% to 85.88%). For cases with US N2–3, the predictive model showed comparable performance to axilla US.

Prediction of residual metastasis by risk stratification system based on dual-modality US

Table 5 displays the pathologic results of axillary LNs after NAT stratified by the simplified risk score calculated form this predictive model based on dual-modality US. Lower risk scores indicate a lower probability of residual axillary LNs metastasis. The axillary non-pCR rate was only 5.21% (5/96) for cases with a risk score of 1, and it increased to 92.59% (75/81) for cases with a risk score of 5. Based on the predictive probability calculated using multivariate LR formula, a risk score of 1 was assigned to cases classified as US N0 with SWV_{mean} <2.77 m/s or as US N1 with SWV_{mean} <1.92 m/s. This risk stratification system was tested using the external test cohort, as shown in Table 6. The axillary non-pCR rate was only 6.97% (4/58) for cases with a risk score of 1, while it was 91.23% (52/57) for cases with a risk score of 5. The effectiveness of the predictive model is illustrated in Supplementary eFigure 1 and eFigure 2 http://links.lww.com/JS9/E116.

Figure 5 and Figure 6 illustrate the clinical benefit of identifying cases scoring 1 as suitable candidates for avoiding ALND. Among patients who underwent ALND in both the training and external test cohort, the rate of unnecessary ALND ranged from 51.79% to 52.53%, and the final benefit rate ranged from 47.47% to 48.21%. Utilizing the risk stratification system based on dual-modality US, patients scoring 1 may safely avoid ALND, significantly reducing unnecessary ALND rate of 17.12%–24.10% and increasing final benefit rate of 73.84%–80.94%. Only 1.95%–2.05% of patients were incorrectly predicted to have axillary pCR by the risk stratification system.

Assessment of risk stratification system stratified across molecular subtype

It is widely recognized that molecular subtype serves as a robust predictor of response to NAT. Therefore, we conducted

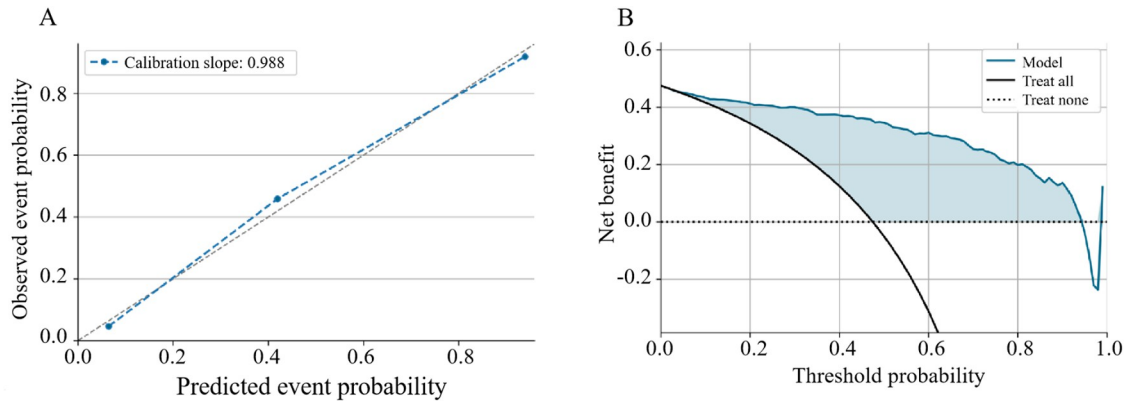

Figure 4. Calibration curve (A) and decision curve (B) for the predictive model based on dual-modality US. Abbreviations: US, ultrasound.

Table 4

The performance of the predictive model based on dual-modality US stratified across US nodal stages (*n* = 257)

US nodal stage	No. of patients	Accuracy (%)	Sensitivity (%)	Specificity (%)	No. of false negative cases	No. of false positive cases
N0	120	88.33 (106/120)	69.23 (18/26)	93.62 (88/94)	8	6
N1	85	85.88 (73/85)	87.76 (43/49)	83.33 (30/36)	6	6
N2	31	87.10 (27/31)	100.00 (27/27)	0.00 (0/4)	0	4
N3	21	95.24 (20/21)	100.00 (20/20)	0.00 (0/1)	0	1

Abbreviation: US, ultrasound.

a subgroup analysis of the risk stratification system based on molecular subtypes. As shown in Supplementary eTable 3, 4, and 5 http://links.lww.com/JS9/E116, the probability of residual axillary LNs metastasis after NAT varied across different molecular subtypes at each risk score. For the HR+HER2– subtype, patients scoring 1 exhibited a higher axillary non-pCR rate of 20.00% (4/20) in the training cohort and 16.67% (2/12) in the external test cohort. For the HR±HER2+ subtype, patients scoring 1 demonstrated 1.54% (1/65) and 2.44% (1/41) in the training and external test cohorts, respectively. For the HR+HER2+ subtype, the two rates were 0.00% (0/11) and 20.00% (1/5).

Discussion

This study demonstrated that, compared with conventional axilla US, breast SWE exhibited superior diagnostic performance for axillary LNs after NAT, particularly in cases with nodal stage of N0–1 observed on axilla US. Breast SWE significantly reduced the numbers of false negatives among cases with US N0 and false positives among cases with US N1. A predictive model incorporating axilla US nodal stage and breast SWE showed good performance for detecting residual axillary LN metastasis. Patients with a risk score of 1 could potentially avoid ALND after NAT, especially those with HR±HER2+ subtype. This model has the potential to assist clinicians in making surgical decisions regarding the axilla, thereby optimizing the overall treatment of patients.

There is currently no consensus on the optimal approach for axillary management after NAT in patients with node-positive breast cancer. The accurate diagnosis of axillary LNs on imaging following NAT is critical to minimize surgical overtreatment of the axilla in breast cancer patients with nodal involvement[29]. US is the primary approach for the non-invasive assessment of residual disease in axillary LNs after NAT[11]. However, previous studies have reported inconsistent and suboptimal performance of conventional axilla US for this application[28,30]. Our findings further demonstrated that false negatives were prevalent in cases with US N0 stage, while false positives were common in cases with US N1 stage. Notably, when axilla US staging indicated N2 or N3 stages, most cases were found to have residual metastasis in axillary LNs. Therefore, it is important to acknowledge that the limitations of axilla US following NAT predominantly affect breast cancer patients with US N0–1 stages.

In this context, our results indicate that breast SWE has promise in diagnosing axillary LNs following NAT in patients with node-positive breast cancer, particularly in those with US N0–1 stages. In recent years, the development of ultrasonic elastography, particularly SWE based on VTIQ, has yielded supplementary data http://links.lww.com/JS9/E116 for LN characterization and has improved the diagnostic accuracy of US in assessing axillary LNs[15-18]. A recent research has demonstrated that SWV value obtained through VITQ can be utilized to characterize pathological features such as tumor cell density and collagen volume fraction following NAT[18]. Furthermore, in comparison with SWE conducted on axillary LNs, breast SWE exhibits superior performance in predicting axillary pCR[18]. However, similar to previous study[18], our results suggested that the overall performance of breast SWE in evaluating axillary LNs after NAT in breast cancer with nodal involvement is still unsatisfactory for clinical application. Considering the varying utility of less radical restaging methods following NAT in patients with different clinic nodal stages[2], we hypothesize that SWE may also exhibit diverse performance in assessing axillary LNs in breast cancer patients with different nodal stages. Our results suggested that breast SWE can overcome the challenges with false negatives for cases with US N0 stage and false positives for cases with US N1 stage after NAT. However, breast SWE did not provide added diagnostic value for assessing axillary LNs in US N2–3 cases.

Table 5

Surgical-pathologic results of axillary LNs after NAT stratified across simplified risk score based on the predictive model in the training cohort (*n* = 257)

Risk score	No. of patients	Axillary pCR, *n*	Axillary non-pCR, *n*	Axillary non-pCR rate (%)
1	96	91	5	5.21
2	32	24	8	25.00
3	24	10	14	58.33
4	24	4	20	83.33
5	81	6	75	92.59

Abbreviations: LN, lymph node; NAT, neoadjuvant therapy; pCR, pathologic complete response.

Table 6

Surgical-pathologic results of axillary LNs after NAT stratified across simplified risk score based on the predictive model in the external test cohort (*n* = 195)

Risk score	No. of patients	Axillary pCR, *n*	Axillary non-pCR, *n*	Axillary non-pCR rate (%)
1	58	54	4	6.97
2	47	3s1	16	34.04
3	14	6	8	57.14
4	19	5	14	73.68
5	57	5	52	91.23

Abbreviations: LN, lymph node; NAT, neoadjuvant therapy; pCR, pathologic complete response.

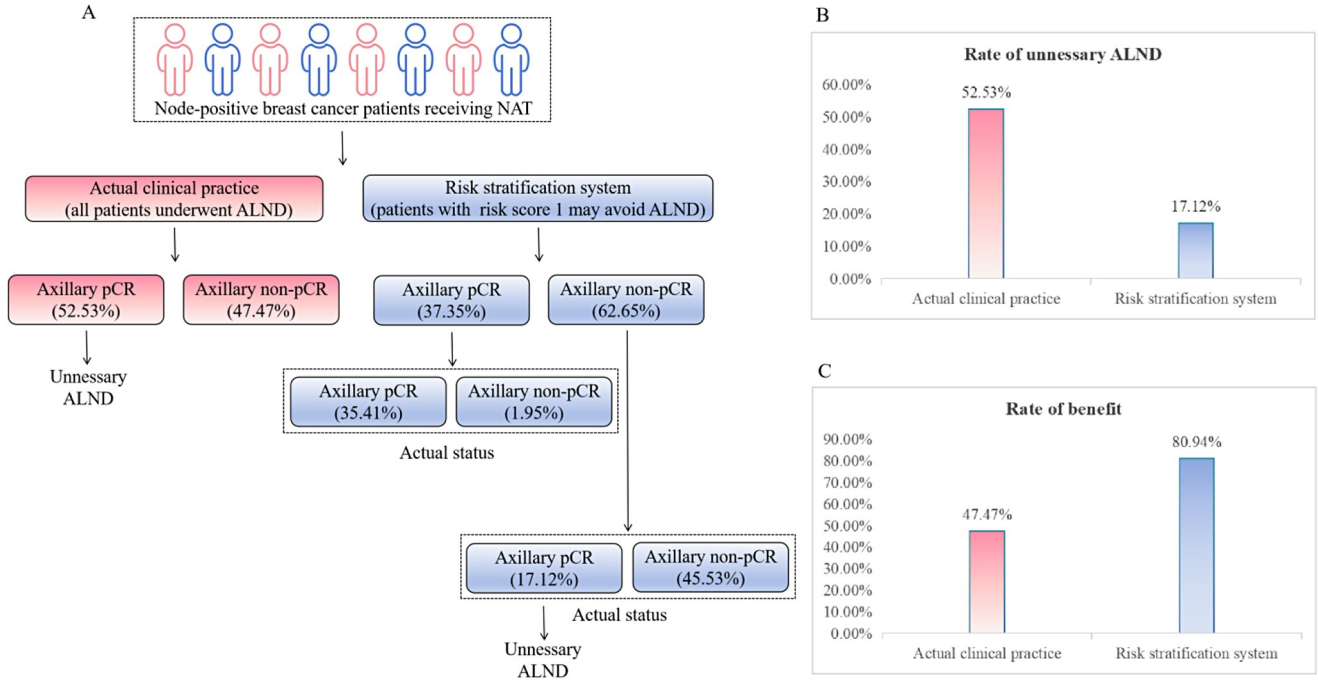

Figure 5. Clinical benefit assessment of risk stratification system based on dual-modality US in the training cohort (*n* = 257). (A) Recommendation for ALND according to risk stratification system based on dual-modality US after NAT in patients with node-positive breast cancer. A risk score of 1 was established as a reliable predictor of axillary pCR. Only 1.95% of patients were incorrectly predicted to have axillary pCR by the risk stratification system. (B) Rate of unnecessary ALND assessed in two scenarios: if all patients underwent ALND, an unnecessary ALND rate of 52.53% was observed; if ALND was performed only in patients with risk score of 2-5, a reduced unnecessary ALND rate of 17.12% was observed. (C) Rate of benefit assessed in two scenarios: if all patients underwent ALND, a final benefit rate of 47.47% was observed; if ALND was performed only in patients with risk score of 2-5, an increased final benefit rate of 80.94% was observed. Abbreviations: US, ultrasound; ALND, axillary lymph node dissection; NAT, neoadjuvant therapy; pCR, pathologic complete response.

In view of the significant differences in axillary pCR rates across different US nodal stages following NAT and the varied performance of breast SWE in patients with varying US nodal stages, we speculated that integrating US nodal stage as a risk factor with breast SWE would provide substantial predictive value. We then developed a predictive model using US nodal stage and breast SWE. The performance of this predictive model (AUC, 0.93) was better than those of models that included US characteristics, with previously reported AUC values between 0.70 and 0.90[27,28,31-34]. Additionally, compared with the false negative rates of SLNB after NAT observed in several clinical trials[2,4], our predictive model based on dual-modality US demonstrated a comparable overall performance, with an overall false negative rate of 11.48%. A false negative rate of less than 10% could be considered acceptable for omitting ALND[2,4,28]. When using this predictive model to calculate the probability of residual metastasis in axillary LNs for developing a risk stratification system, the rate of axillary non-pCR was only 5.21%–6.97% in cases with a risk score of 1, indicating a reduction in false negative results if cases scoring 1 are identified as axillary pCR[28]. The risk stratification system based on dual-modality US provides clinicians with more detailed references regarding the selection of axilla management for specific individuals after NAT. Patients with a risk score of 1 (US N0 with SWV_{mean} <2.77 m/s or US N1 with SWV_{mean} <1.92 m/s) may be appropriate candidates for less radical management to avoid unnecessary ALND and its complications. The reduction in unnecessary ALND rate and the increase in benefit rate demonstrate its clinical applicability. Furthermore, only 1.95%–2.05% of patients were incorrectly predicted to have axillary pCR, indicating the safety of this risk stratification system.

While the developed risk stratification system demonstrated excellent overall performance in evaluating residual axillary LNs metastasis after NAT, varying risk probabilities were observed across different molecular subtypes. Patients with HR+HER2– breast cancer showed a higher axillary non-pCR rate compared to other subtypes. Surprisingly, even patients scoring 1 had a high non-pCR rate between 16.67% and 20.00%. This could be attributed to the relative insensitivity of the HR+HER2– subtype to NAT. In contrast, patients with the HR±HER2+ subtype scoring 1 indicated a low axillary non-pCR rate from 1.54% to 2.44%. Therefore, patients with HR±HER2+ scoring 1 might be better candidates for axilla conservation post-NAT. In cases of HR- HER2- breast cancer, the axillary non-pCR rates could not be robustly assessed due to the limited sample size.

In this multicenter, prospective, observational study, the predictive model incorporating both axilla US nodal stage and breast SWE exhibited excellent predictive ability. A simplified risk score generated from this predictive model could help personalize treatment strategies and reduce unnecessary interventions. This model serves as a practical and convenient tool for clinicians, as the variables it uses are readily accessible and do not require additional interventional procedures. Furthermore, decision curve analysis indicated that the model provides satisfactory clinical utility in facilitating individualized strategies for

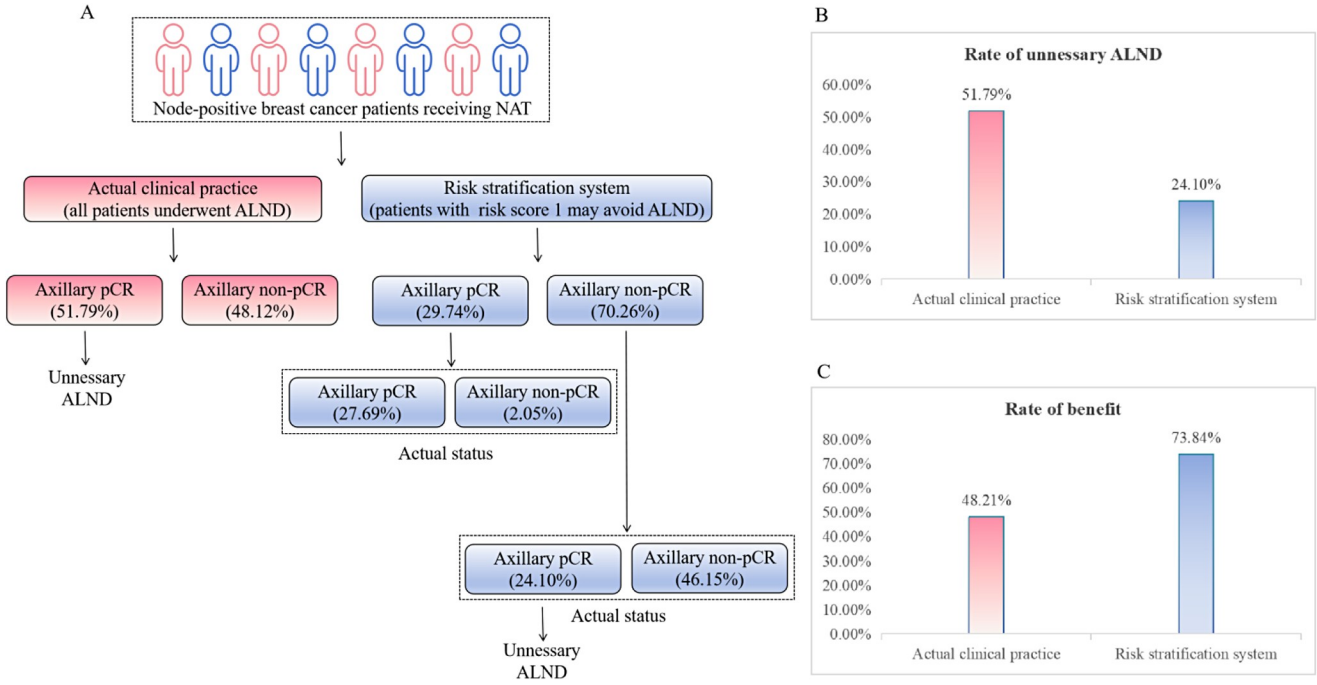

Figure 6. Clinical benefit assessment of risk stratification system based on dual-modality US in the external test cohort (n = 195). (A) Recommendation for ALND according to risk stratification system based on dual-modality US after NAT in patients with node-positive breast cancer. A risk score of 1 was established as a reliable predictor of axillary pCR. Only 2.05% of patients were incorrectly predicted to have axillary pCR by the risk stratification system. (B) Rate of unnecessary ALND assessed in two scenarios: if all patients underwent ALND, an unnecessary ALND rate of 51.79% was observed; if ALND was performed only in patients with risk score of 2-5, a reduced unnecessary ALND rate of 24.10% was observed. (C) Rate of benefit assessed in two scenarios: if all patients underwent ALND, a final benefit rate of 48.21% was observed; if ALND was performed only in patients with risk score of 2-5, an increased final benefit rate of 73.84% was observed. Abbreviations: US, ultrasound; ALND, axillary lymph node dissection; NAT, neoadjuvant therapy; pCR, pathologic complete response.

axilla LNs after NAT in breast cancer patients with nodal involvement. Despite encouraging findings, this study has several limitations. First, both SWE and US have limitations in terms of depth of penetration, which may impact their ability to accurately assess deeper LNs, potentially missing important information. Second, SWE may not be universally available in all healthcare settings due to cost constraints or limited access to specialized equipment, which could limit their widespread applicability. Third, the variability in SWE measurements from different vendors remains a problem that needs to be addressed in future studies. Fourth, while longitudinal imaging data during NAT could enhance the predictive performance for nodal response[31,35], SWE data prior to treatment were not obtained in this study. Finally, the pathological findings from the ALND were used as the gold standard in this study, ensuring that the results were based on definitive histopathological evidence and providing a reliable benchmark for outcome assessment. However, several patients who underwent less radical surgical procedures were excluded, potentially introducing selection bias. Given that SLNB with dual mapping and TAD may offer greater clinical benefits, patients undergoing these procedures will be included in future studies to enhance the clinical relevance and generalizability of the findings.

Conclusion

Breast SWE is effective in assessing axillary LNs after NAT in patients with node-positive breast cancer, especially for those with US N0–1 stage. A predictive model incorporating axilla US nodal stage and breast SWE showed good performance for identifying residual axillary LN metastasis. The risk stratification system derived from this predictive model was capable of estimating the probability of residual metastasis, thereby serving as a valuable reference for individualized management strategies. Patients with a risk score of 1 could potentially benefit from a less aggressive axillary management approach, particularly for those diagnosed with HR±HER2+ breast cancer. Future studies with a larger sample size and more external test sets would provide more convincing evidence for clinical application of this model.

Ethical approval

The study design and protocol were approved by the ethics committee of the institutional review board of Sun Yat-sen University Cancer Center (B2022-373-X01).

Consent

Written informed consent for study participation was obtained from all patients.

Sources of funding

This work was supported in part by the China Department of Science and Technology under Key Grant 2023YFE0204300, in

part by the RD Project of Pazhou Lab (HuangPu) under Grant 2023K0606, in part by the NSFC under Grant 82441027, Grant 62371476, in part by Guangzhou Science and Technology Planning Project under Grant 1563000308, in part by Fostering Program for NSFC Young Applicants (Tulip Talent Training Program) of Sun Yat-sen University Cancer Center under Grant 2025yfd19.

Author contributions

X.Q.P. and J.X.H. conceived and designed the project. J.X.H., J.S.M., F.C., Y.W.W., X.Y.W., and F.T.L. collected and interpreted radiological data. Y.T.T., S.D.Q., C.G.S., G.L.H., and Y.T.Z. provided the clinical and pathological data. J.X.H., J.H.H., and Y.L. provided statistical analysis and figures preparation. J.X.H. drafted the manuscript. X.Q.P. and M.S.C. critically reviewed the manuscript for important intellectual content. All authors read and approved the final manuscript.

Conflict of interest disclosure

The authors declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Guarantor

Xiao-Qing Pei.

Research registration unique identifying number (UIN)

This study was registered on the Chinese Clinical Trial Registry with the registration number ChiCTR2400085035. Additional information about the study can be accessed on the website address http://www.chictr.org/cn/.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Data availability statement

The authors declare that they had full access to all of the data in this study and the authors take complete responsibility for the integrity of the data and the accuracy of the data analysis. Due to the privacy of patients, the data related to patients cannot be available for public access but can be obtained from the corresponding author (peixq@sysucc.org.cn) on reasonable request approved by the institutional review board.

References

[1] Gradishar WJ, Moran MS, Abraham J, *et al*. NCCN guidelines® insights: breast cancer, version 4.2023. J Natl Compr Canc Netw 2023;21:594–608.

[2] Boughey JC, Suman VJ, Mittendorf EA, *et al*. Sentinel lymph node surgery after neoadjuvant chemotherapy in patients with node-positive breast cancer: the ACOSOG Z1071 (Alliance) clinical trial. JAMA 2013;310:1455–61.

[3] Galimberti V, Cole BF, Viale G, *et al*. Axillary dissection versus no axillary dissection in patients with breast cancer and sentinel-node micrometastases (IBCSG 23-01): 10-year follow-up of a randomised, controlled phase 3 trial. Lancet Oncol 2018;19:1385–93.

[4] Kuehn T, Bauerfeind I, Fehm T, *et al*. Sentinel-lymph-node biopsy in patients with breast cancer before and after neoadjuvant chemotherapy (SENTINA): a prospective, multicentre cohort study. Lancet Oncol 2013;14:609–18.

[5] Cao S, Liu X, Cui J, *et al*. Feasibility and reliability of sentinel lymph node biopsy after neoadjuvant chemotherapy in breast cancer patients with positive axillary nodes at initial diagnosis: an up-to-date meta-analysis of 3,578 patients. Breast 2021;59:256–69.

[6] Kirkilesis G, Constantinidou A, Kontos M. False negativity of targeted axillary dissection in breast cancer. Breast Care (Basel) 2021;16:532–38.

[7] Munck F, Jepsen P, Zeuthen P, *et al*. Comparing methods for targeted axillary dissection in breast cancer patients: a nationwide, retrospective study. Ann Surg Oncol 2023;30:6361–69.

[8] de Wild SR, Koppert LB, van Nijnatten TJA, *et al*. Systematic review of targeted axillary dissection in node-positive breast cancer treated with neoadjuvant systemic therapy: variation in type of marker and timing of placement. Br J Surg 2024;111:znae071.

[9] Gasparri ML, de Boniface J, Poortmans P, *et al*. Axillary surgery after neoadjuvant therapy in initially node-positive breast cancer: international EUBREAST survey. Br J Surg 2022;109:857–63.

[10] Simons JM, van Nijnatten TJA, van der Pol CC, *et al*. Diagnostic accuracy of radioactive iodine seed placement in the axilla with sentinel lymph node biopsy after neoadjuvant chemotherapy in node-positive breast cancer. JAMA Surg 2022;157:991–99.

[11] Hayward JH, Linden OE, Lewin AA, *et al*. Expert Panel on Breast Imaging. ACR appropriateness criteria® monitoring response to neoadjuvant systemic therapy for breast cancer: 2022 update. J Am Coll Radiol 2023;20:S125–S145.

[12] Huang JX, Wu L, Wang XY, *et al*. Delta radiomics based on longitudinal dual-modal ultrasound can early predict response to neoadjuvant chemotherapy in breast cancer patients. Acad Radiol 2024;31:1738–47.

[13] Gu J, Zhong X, Fang C, *et al*. Deep learning of multimodal ultrasound: stratifying the response to neoadjuvant chemotherapy in breast cancer before treatment. Oncologist 2024;29:e187–e197.

[14] Huang JX, Shi J, Ding SS, *et al*. Deep learning model based on dual-modal ultrasound and molecular data for predicting response to neoadjuvant chemotherapy in breast cancer. Acad Radiol 2023;30 Suppl 2:S50–S61.

[15] Zheng X, Yao Z, Huang Y, *et al*. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nat Commun 2020;11:1236.

[16] Jiang M, Li CL, Luo XM, *et al*. Radiomics model based on shear-wave elastography in the assessment of axillary lymph node status in early-stage breast cancer. Eur Radiol 2022;32:2313–25.

[17] Kim K, Shim SR, Kim SJ. Diagnostic values of 8 different imaging modalities for preoperative detection of axillary lymph node metastasis of breast cancer: a Bayesian network meta-analysis. Am J Clin Oncol 2021;44:331–39.

[18] Huang JX, Liu FT, Sun L, *et al*. Comparing shear wave elastography of breast tumors and axillary nodes in the axillary assessment after neoadjuvant chemotherapy in patients with node-positive breast cancer. Radiol Med 2024;129:1143–55.

[19] Huang JX, Lu Y, Tan YT, *et al*. Elastography-based AI model can predict axillary status after neoadjuvant chemotherapy in breast cancer with nodal involvement: a prospective, multicenter, diagnostic study. Int J Surg 2025;111:221–9.

[20] Cosgrove DO, Berg WA, Doré CJ, *et al*. BE1 Study Group. Shear wave elastography for breast masses is highly reproducible. Eur Radiol 2012;22:1023–32.

[21] Bossuyt PM, Reitsma JB, Bruns DE, *et al*. STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Clin Chem 2015;61:1446–52.

[22] Fischerova D, Garganese G, Reina H, *et al*. Terms, definitions and measurements to describe sonographic features of lymph nodes: consensus opinion from the Vulvar International Tumor Analysis (VITA) group. Ultrasound Obstet Gynecol 2021;57:861–79.

[23] Gabriel NH, James LC, Carl JD, *et al*. Breast. In: Mahul BA, ed. American Joint Committee on Cancer (AJCC). AJCC Cancer Staging Manual. 8th. New York, NY: Springer; 2017. 589–628.

[24] Huang Y, Liu Y, Wang Y, *et al*. Quantitative analysis of shear wave elastic heterogeneity for prediction of lymphovascular invasion in breast cancer. Br J Radiol 2021;94:20210682.
[25] Huang JX, Liu F, Tan YT, *et al*. Enhancing detection of high-level axillary lymph node metastasis after neoadjuvant therapy in breast cancer patients with nodal involvement: a combined approach of axilla ultrasound and breast elastography. Radiol Med 2025;130: 121–131.
[26] Goldhirsch A, Winer EP, Coates AS, *et al*. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2013. Ann Oncol 2013;24:2206–23.
[27] You J, Huang Y, Ouyang L, *et al*. Automated and reusable deep learning (AutoRDL) framework for predicting response to neoadjuvant chemotherapy and axillary lymph node metastasis in breast cancer using ultrasound images: a retrospective, multicentre study. EClinicalMedicine 2024;69:102499.
[28] Kim R, Chang JM, Lee HB, *et al*. Predicting axillary response to neoadjuvant chemotherapy: breast MRI and US in patients with node-positive breast cancer. Radiology 2019;293:49–57.
[29] Boughey JC, Ballman KV, Hunt KK, *et al*. Axillary ultrasound after neoadjuvant chemotherapy and its impact on sentinel lymph node surgery: results from the American college of surgeons oncology group Z1071 trial (Alliance). J Clin Oncol 2015;33:3386–93.
[30] Banys-Paluchowski M, Gruber IV, Hartkopf A, *et al*. Axillary ultrasound for prediction of response to neoadjuvant therapy in the context of surgical strategies to axillary dissection in primary breast cancer: a systematic review of the current literature. Arch Gynecol Obstet 2020;301:341–53.
[31] Fu Y, Lei YT, Huang YH, *et al*. Longitudinal ultrasound-based AI model predicts axillary lymph node response to neoadjuvant chemotherapy in breast cancer: a multicenter study. Eur Radiol 2024;34:7080–89.
[32] Gu J, Tong T, Xu D, *et al*. Deep learning radiomics of ultrasonography for comprehensively predicting tumor and axillary lymph node status after neoadjuvant chemotherapy in breast cancer patients: a multicenter study. Cancer 2023;129:356–66.
[33] Huang JX, Chen YJ, Wang XY, *et al*. Nomogram based on US and clinicopathologic characteristics: axillary nodal evaluation following neoadjuvant chemotherapy in patients with node-positive breast cancer. Clin Breast Cancer 2024;24:e452–e463.e4.
[34] Shi W, Huang X, Wang Y, *et al*. A novel nomogram containing efficacy indicators to predict axillary pathologic complete response after neoadjuvant systemic therapy in breast cancer. Front Endocrinol (Lausanne) 2022;13:1042394.
[35] Liu Y, Wang Y, Wang Y, *et al*. Early prediction of treatment response to neoadjuvant chemotherapy based on longitudinal ultrasound images of HER2-positive breast cancer patients by Siamese multi-task network: a multicentre, retrospective cohort study. EClinicalMedicine 2022;52: 101562.