

Risk factors for axillary lymph node metastases in clinical stage T1-2N0M0 breast cancer patients

Yuanxin Zhang, MD^a, Ji Li, MD^{a,b}, Yuan Fan, MD^a, Xiaomin Li, MD^a, Juanjuan Qiu, MD^a, Mou Zhu, MD^c, Hongjiang Li, MD^{a,*}

Abstract

Axillary lymph node metastasis (ALNM) is commonly the earliest detectable clinical manifestation of breast cancer when distant metastasis emerges. This study aimed to explore the influencing factors of ALNM and develop models that can predict its occurrence preoperatively.

Cases of sonographically visible clinical stage T1-2N0M0 breast cancers treated with breast and axillary surgery at West China Hospital were retrospectively reviewed. Univariate and multivariate logistic regression analyses were performed to evaluate associations between ALNM and variables. Decision tree analyses were performed to construct predictive models using the C5.0 packages.

Of the 1671 tumors, 541 (32.9%) showed axillary lymph node positivity on final surgical histopathologic analysis. In multivariate logistic regression analysis, tumor size ($P < .001$), infiltration of subcutaneous adipose tissue ($P < .001$), infiltration of the interstitial adipose tissue ($P = .031$), and tumor quadrant locations ($P < .001$) were significantly correlated with ALNM. Furthermore, the accuracy in the decision tree model was 69.52%, and the false-negative rate (FNR) was 74.18%. By using the error-cost matrix algorithm, the FNR significantly decreased to 14.75%, particularly for nodes 5, 8, and 13 (FNR: 11.4%, 9.09%, and 14.29% in the training set and 18.1%, 14.71%, and 20% in the test set, respectively).

In summary, our study demonstrated that tumor lesion boundary, tumor size, and tumor quadrant locations were the most important factors affecting ALNM in cT1-2N0M0 stage breast cancer. The decision tree built using these variables reached a slightly higher FNR than sentinel lymph node dissection in predicting ALNM in some selected patients.

Abbreviations: ALND = axillary lymph node dissection, ALNM = axillary lymph node metastasis, ALNS = axillary lymph node status, BMI = body mass index, FNR = false-negative rate, FPR = false-positive rate, IBC = invasive breast cancer, IG = information gain, IIAT = infiltration of the interstitial adipose tissue, ISAT = infiltration of subcutaneous adipose tissue, NIIAT = no infiltration of the interstitial adipose tissue, NISAT = no infiltration of subcutaneous adipose tissue, non-ALNM = no axillary lymph node metastasis, OS = overall survival, SLN = sentinel lymph node, SLNB = sentinel lymph node biopsy, SLNM = sentinel lymph node metastases, TQLS = tumor quadrant locations, US = ultrasonography.

Keywords: axillary lymph node metastasis, breast cancer, decision tree, risk factor

1. Introduction

Metastatic spread from primary breast cancer can occur during the early stage, and axillary lymph node metastasis (ALNM) is

usually the earliest detectable clinical presentation when distant metastasis emerges.^[1] Sentinel lymph node biopsy (SLNB) is the standard approach for axillary staging in breast cancer patients with no palpable axillary adenopathy, and the trend for breast cancer treatment is currently shifting towards minimizing axillary surgery, even in the presence of sentinel lymph node (SLN) involvement. The Z0011 trial has demonstrated that the 10-year overall survival (OS) of patients with 1 to 2 sentinel lymph node metastases (SLNM) treated with SLNB and whole-breast irradiation is non-inferior to the OS of patients with no palpable axillary adenopathy clinical T1-T2 invasive breast cancer (IBC) treated with axillary lymph node dissection (ALND).^[2] Moreover, the AMAROS trial found further evidence to show that the axillary recurrence-free rate of patients with SLNB-proven metastasis treated with radiotherapy is non-inferior to those treated with ALND. The radiotherapy group not only had lower lymphedema rate at 1 or 5 years during treatment, but also developed fewer surgery-related complications.^[3] Collectively, the results of these trials suggest that IBC is a systemic disease,^[4,5] and as such, the treatment strategy should start with identification of the lymph node status and complete assessment of the TNM stage more than just immediate treatment.^[6,7]

Although SLNB remains the standard of care for breast cancer patients with no palpable axillary adenopathy, approximately 24.8% to 35.5% have ALNM in the final pathological

Editor: Chinnadurai Mani.

YZ and JL contributed equally to the study.

This study was supported by a grant from the Science and Technology Department of Sichuan Province (2017SZ0065).

The authors have no conflicts of interest to disclose.

^a Department of Breast Surgery, ^b Anesthesia surgery center, ^c Department of Ultrasound, West China Hospital, Sichuan University, Chengdu, China.

* Correspondence: Hongjiang Li, Department of Breast Surgery, West China Hospital, Sichuan University, 37 Guoxue Street, Chengdu 610041, China (e-mail: lihongjiang@sohu.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

How to cite this article: Zhang Y, Li J, Fan Y, Li X, Qiu J, Zhu M, Li H. Risk factors for axillary lymph node metastases in clinical stage T1-2N0M0 breast cancer patients. *Medicine* 2019;98:40(e17481).

Received: 26 February 2019 / Received in final form: 16 July 2019 / Accepted: 12 September 2019

<http://dx.doi.org/10.1097/MD.00000000000017481>

results,^[8–11] showing that a high number of patients receive excessive medical care. Consequently, noninvasive methods for assessing axillary lymph nodes is urgently needed. The ongoing SOUND trial, which recruits patients with small tumors (≤ 2 cm) and negative lymph node on ultrasonography (US), is designed to confirm the diagnostic efficiency of US, but the results are yet to be determined.^[12]

Previous studies have shown that clinicopathological features such as size, age at diagnosis, palpable mass or not, body mass index (BMI) and hormone receptors and so on are related to ALNM.^[13–16] Meanwhile, high-frequency ultrasonography, the first-line imaging modality in breast cancer diagnosis, can show the rich morphological features of breast tumor, and some of those features may be related to ALNM.^[13,17–20]

Therefore, this study aimed to explore the influencing factors of ALNM with respect to both ultrasonographic characteristics and clinicopathologic traits. The total population was divided into the training set and test set. We used the machine learning method to establish a predictive model for cT1-T2N0M0 patients, and its accuracy was evaluated to provide a preliminary experimental basis for clinical research and related treatment. Clinicians may reference the prediction results and make better clinical decisions.

2. Materials and methods

2.1. Study participants

A consecutive cohort of patients who underwent curative-intent breast surgery were reviewed from January 1, 2014 to December

31, 2017. The eligibility criteria were (1) female sex; (2) cT1-T2N0M0 stage; (3) visible tumors on pre-biopsy US; (4) complete clinicopathological and ultrasonographic data; and (5) surgery within the next 2 months since US. These criteria were set based on the fact that biopsy may change the morphological features of the tumor and that undergoing surgery within 2 months after US can avoid the risk of an altered axillary status. Meanwhile, the exclusion criteria are shown in Figure 1.

This study was approved by the Institutional Review Board of West China Hospital and performed in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from all patients.

2.2. Ultrasound imaging and analysis

Breast and axilla US were performed using a linear array probe (5–15 MHz) supplemented by the 1 to 5 MHz convex array probe as needed to penetrate larger masses (Philips iU22 and HDI 5000, Philips Medical Systems, Bothell, WA, USA; HI VISION Preirus, Hitachi Medical, Tokyo, Japan; Esaote MyLab 90, Esaote, Genova, Italy; GE Logiq E9, General Electric Healthcare, Milwaukee, WI, USA). All US exams were performed by experienced sonographers. The US findings were retrospectively analyzed by an experienced sonographer based on the criteria from the ACR BI-RADS lexicon for US. Lesion boundary (LB) is not a BI-RADS term but is used in our Ultrasound department.^[21,22] According to the relationship between the mass and

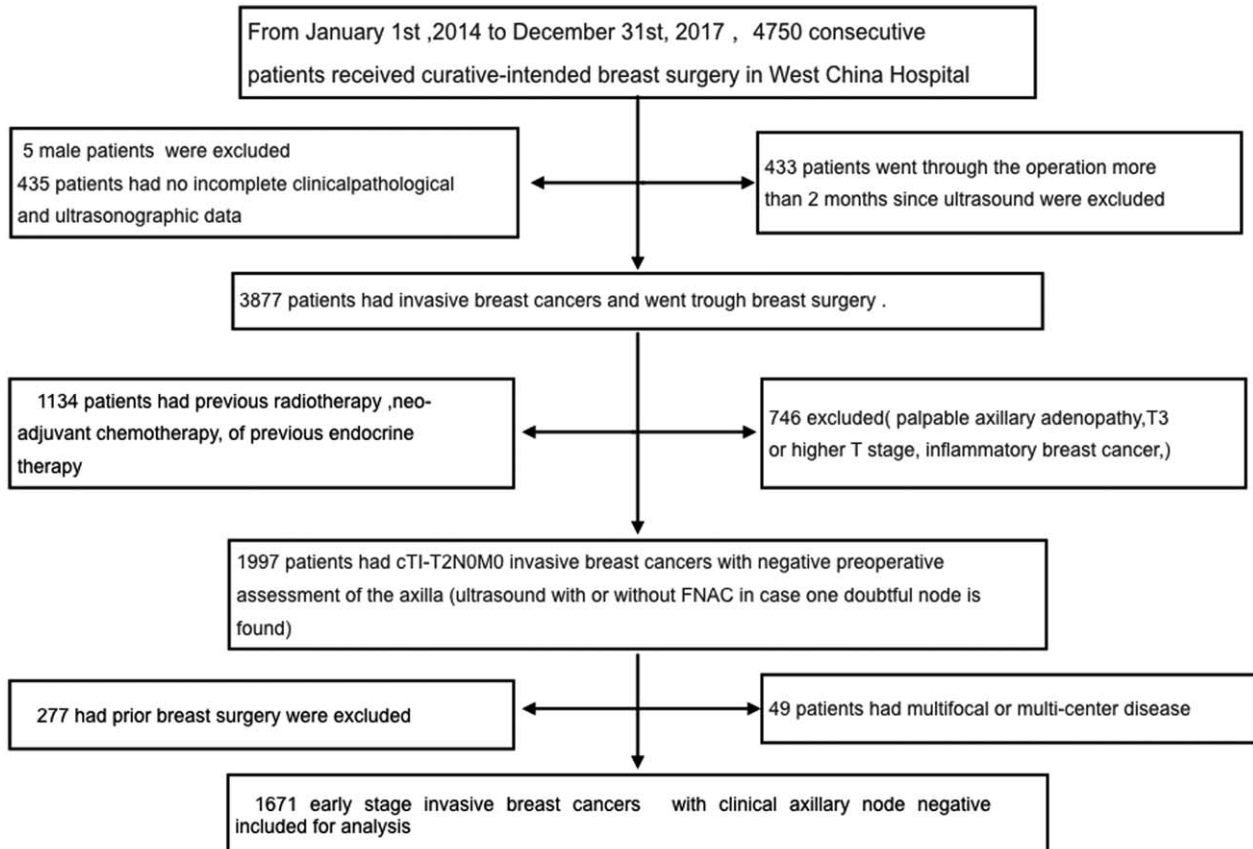


Figure 1. Inclusion and exclusion flow diagram.

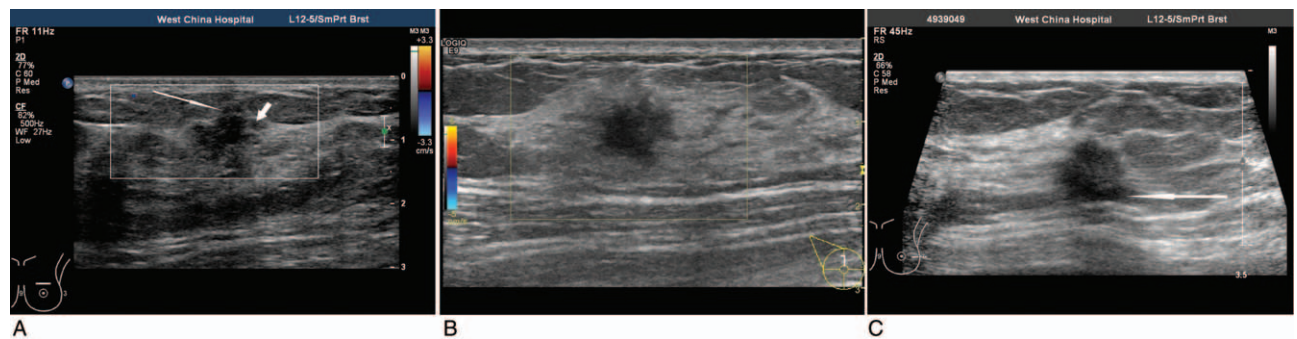


Figure 2. Ultrasound images showing different breast tumor lesion boundary. a: Ultrasound image of a 13*6*11 mm IBC in a 46-year-old woman shows that the tumor infiltrates subcutaneous adipose tissue (long thin arrow) with an angular margin (short thin arrow). The final pathological diagnosis after ALND confirmed ALNM (5/23). b: A 11*9*10 mm IBC with a hypoechoic mass that was confined to the mammary gland, with indistinct margins in a 49-year-old woman. 0/3 sentinel lymph node macrometastasis (SLNM). c: Ultrasound image of a 11*11*12 mm IBC in a 55-year-old woman showing tumor infiltration of the interstitial adipose tissue (long thick arrow), 2/3 sentinel lymph node (SLN) macrometastasis, and 4/24 ALNM after ALND.

the breast tissue in US, LB on US were divided into infiltration of subcutaneous adipose tissue (ISAT), infiltration of the interstitial adipose tissue (IIAT), and confined to the gland among cT1-T2N0M0 patients. Cases where there was an obscure or fuzzy boundary that made it difficult to identify whether the tumor extended to the subcutaneous or interstitial adipose tissue and were enveloped or there were tissue overlaps that caused artifacts on imaging were defined as “obscure boundary”.^[17,23] Other US parameters such as tumor size, tumor quadrant locations (TQLs), and tumor distance from the nipple were collected. Samples of different LB are shown in Figure 2a–c.

2.3. Decision tree prediction

C5.0 decision trees, a machine learning classification algorithm implemented in the R environment, were used to construct a statistical classifier.^[24,25] C5.0 decision trees depend on the concepts of information gain (IG) and entropy to determine the attributes that provide the highest information about the instances on which the tree is modeled. The IG is calculated for each of the attributes and the 1 with the smallest entropy (highest IG) forms the root node of the tree.^[26] Given that the decision tree tends to overfit the training data model, the reported error rate in the training data may be too optimistic. Therefore, it is important to evaluate the decision tree model based on the test data set. Our study used “cross validation” method to divide data and obtained 900 for training set and 761 for test set.

2.4. Statistical analyses

All statistical analyses were performed using SPSS (version 25.0, IBM Corporation). Group difference for continuous and categorical variables was assessed using Student *t* test and normal (Pearson) Chi-Squared test, respectively. Univariate and multivariate logistic regression analyses were performed using a stepwise selection of all features studied as candidate predictors of axillary lymph node status. All tests were two-sided, and a *P* value < .05 was considered statistically significant. For independent factors identified via multivariate logistic regression analyses, decision tree analyses were performed to construct a decision tree model using the C5.0 packages within the R environment.

3. Results

3.1. Clinicopathological characteristics

During the study period, 4750 patients were diagnosed with operable IBC. Of these, 3079 were excluded, and 1671 were evaluated. The mean age was 51.27 ± 11.7 years, and the average body mass index was in the normal range (23.08 ± 2.94). More than 50% of the patients were on menopause (51.53%) at the time of cancer diagnosis, and a large proportion of tumors were palpable (81.33%) or located in the lateral quadrant (46.2%). In total, most cases (1469 cases (88%)) of breast cancers were invasive ductal carcinoma. The majority of IBC were confined in the gland, and the tumor rarely infiltrated both subcutaneous and interstitial adipose tissue (Table 1).

In total, 541 (32.9%) patients were diagnosed with ALNM breast cancer on immunohistochemistry analysis of the surgical specimen. In correlation analysis, axillary lymph node status (ALNS) was significantly correlated with tumor size in US ($P < .001$), tumor distance from the nipple ($P = .017$), palpability ($P < .001$), TQLs ($P < .001$), histologic grade ($P = .01$), ISAT ($P < .001$), and IIAT ($P < .001$). However, menopausal status, molecular subtypes, and tumor pathological subtypes were not significantly different between ALNM patients and non-ALNM patients (Table 1).

3.2. Univariate and multivariate logistic regression analysis

To identify the clinical traits affecting ALNM, additional univariate logistic regression was performed on all candidate predictors. The results showed that tumor size (OR=1.052, $P < .001$), tumor distance from the nipple (OR=0.94, $P = .018$), palpability (OR=2.200, $P < .001$), TQLs ($P < .001$), histologic grade ($P < .001$), ISAT ($P < .001$), and IIAT ($P = .003$) were significantly correlated with ALNS (Table 2). Features statistically significant in the univariate logistic regression model were included in the multivariable logistic regression model. The results showed that size, TQLs, and LB were independent risk factors of ALNM, indicating that larger tumor size, tumor location beneath the nipple, ISAT, and IIAT are high-risk factors for ALNM (Table 2).

3.3. Decision tree prediction

Independent factors identified via multivariate logistic regression analysis including tumor size, TQLs, ISAT, and IIAT were

Table 1**Clinicopathologic characteristics of the Cohort.**

Variables	All (N = 1671)	ALNM (N = 1130)	Non-ALNM (N = 541)	t	χ^2	P value
Age (years)	51.27 ± 11.7	51.44 ± 11.54	50.92 ± 12.03	0.85		.395
Size (mm) [§]	22.35 ± 9.11	21.00 ± 8.67	25.18 ± 9.35	-9.00		<.001
Proximity to the nipple (cm)	3.80 ± 2.15	3.89 ± 2.11	3.62 ± 2.20	2.38		.017
BMI					2.46	.117
≤24	1130	716	364			
>24	541	414	177			
Palpable					27.40	<.001
Yes	1359	880	479			
No	312	250	62			
Pain					2.34	.126
Yes	403	260	143			
No	1268	870	398			
TQLs					40.19	<.001
Medial	210	152	58			
Overlapping*	326	258	68			
Beneath nipple	363	208	155			
Lateral	772	512	260			
Menopausal status					2.27	.132
Premenopausal	808	532	276			
Postmenopausal	863	598	265			
T grade					52.20	<.001
T1 (≤2CM)	775	593	182			
T2 (>2, ≤5CM)	896	537	359			
ER					1.56	.212
Negative	399	280	119			
Positive	1272	850	422			
PR					1.39	.238
Negative	451	315	136			
Positive	1220	815	405			
Her2					1.49	.476
Negative	1027	699	328			
Positive	444	291	153			
Uncertain	200	140	60			
Ki67					0.93	.336
Negative (<20)	379	264	115			
Positive (≥20)	1292	866	426			
Pathological subtypes					0.54	.765
IDC	1469	992	477			
ILC	102	72	30			
Others	100	66	34			
Histologic grade					11.27	.01
1	212	163	49			
2	726	482	244			
3	682	447	235			
Uncertain	51	38	13			
Molecular subtypes					2.76	.599
Luminal A	144	98	46			
Luminal B	1047	695	352			
Her2-enriched	122	83	39			
Basal-like	158	114	44			
Uncertain	200	140	60			
ISAT					91.73	<.001
No	1221	906	315			
Uncertain	29	18	11			
Yes	421	206	215			
IIAT					11.56	.003
No	1161	815	346			
Uncertain	337	207	130			
Yes	173	108	65			

* Overlapping means the tumor was located in the junction between the medial and lateral position (at the 6 o'clock and 12 o'clock position).

§Tumor size means the tumor's longest length in ultrasound images.

ALNM = axillary lymph node metastasis, ER = estrogen receptor, Her2 = human epidermal growth factor receptor 2, IDC = invasive ductal carcinoma, IIAT = infiltration of the interstitial adipose tissue, ILC = invasive lobular carcinoma, ISAT = infiltration of subcutaneous adipose tissue, non-ALNM = no axillary lymph node metastasis, PR = progesterone receptor, T grade = tumor size grade, TQLs = tumor quadrant locations.

Table 2**Univariate and multivariate logistic regression analysis.**

Variables	Univariate analysis		Multivariate analysis	
	Odds ratio	P value	Odds ratio	P value
Size (mm) ^S	1.05 (1.04,1.06)	<.001	1.048 (1.036,1.061)	<.001
Proximity to the nipple (cm)	0.94 (0.90,0.99)	.018	1.049 (0.979,1.124)	.175
BMI		.12		
<=24	1			
>24	1.19 (0.96,1.48)	.12		
Palpable		<.001		.598
No	1		1	
Yes	2.20 (1.63,2.96)	<.001	1.107 (0.758,1.617)	.598
TQLs		<.001		<.001
Medial	1		1	
Overlapping* quadrants (6,12)	0.70 (0.46,1.03)	.072	0.684 (0.449,1.042)	.077
Beneath nipple	1.95 (1.35,2.82)	<.001	1.694 (1.152,2.491)	.007
Lateral	1.33 (0.95,1.87)	.097	1.250 (0.879,1.779)	.214
T grade		<.001		.646
T1	1		1	
T2	2.18 (1.76,2.70)	<.001	1.087 (0.761,1.553)	.646
ER		.212		
Negative	1			
Positive	0.86 (0.67,1.09)	.212		
PR		.24		
Negative	1			
Positive	0.87 (0.69,1.10)	.24		
Her2		.48		
Negative	1			
Positive	1.12 (0.89,1.42)	.34		
Uncertain	0.91 (0.66,1.27)	.59		
KI67		.34		
Negative (<20)	1			
Positive (>=20)	1.13 (0.88,1.45)	.34		
Pathological subtype		.765		
IDC	1			
ILC	0.87 (0.56,1.35)	.52		
Others	1.07 (0.70,1.64)	.75		
Histologic grade		.011		.128
1	1		1	
2	1.68 (1.18,2.40)	.004	1.465 (1.010,2.124)	.044
3	1.75 (1.23,2.50)	.002	1.544 (1.062,2.243)	.023
Uncertain	1.14 (0.56,2.31)	.72	1.154 (0.549,2.427)	.705
Molecular subtypes		.601		
Luminal A	1			
Luminal B	1.08 (0.74,1.57)	.69		
Her2-enriched	1.00 (0.60,1.68)	1.00		
Basal-like	0.82 (0.50,1.35)	.44		
Uncertain	0.91 (0.58,1.45)	.70		
ISAT		<.001		<.001
No	1		1	
Uncertain	1.76 (0.82,3.76)	.146	1.361 (0.615,3.012)	.446
Yes	3.00 (2.39,3.78)	<.001	2.717 (2.14,3.451)	<.001
IIAT		.003		.031
No	1		1	
Uncertain	1.48 (1.15,1.91)	.002	1.408 (1.077,1.842)	.012
Yes	1.42 (1.02,1.98)	.04	1.261 (0.882,1.803)	.203

* Overlapping means the tumor was located in the junction between the medial and lateral position (at the 6 o'clock and 12 o'clock position).

^STumor size means the tumor's longest length in ultrasound images.

ALNM = axillary lymph node metastasis, ER = estrogen receptor, Her2 = human epidermal growth factor receptor 2, IDC = invasive ductal carcinoma, IIAT = infiltration of the interstitial adipose tissue, ILC = invasive lobular carcinoma, ISAT = infiltration of subcutaneous adipose tissue, non-ALNM = no axillary lymph node metastasis, PR = progesterone receptor, T grade = tumor size grade, TQLs = tumor quadrant locations.

selected to build the decision tree. Among the 1671 patients, we assigned 900 (54%) patients in the training sample, and the remaining 771 (46%) patients were enrolled in the test sample by "cross validation" method. As shown in Figure 3a, high

consistency was obtained in the training set and test set after machine learning. Therefore, we believe that both sets can adequately represent the entire research cohort. In the accuracy model, the decision tree was developed using ISAT (100.00%),

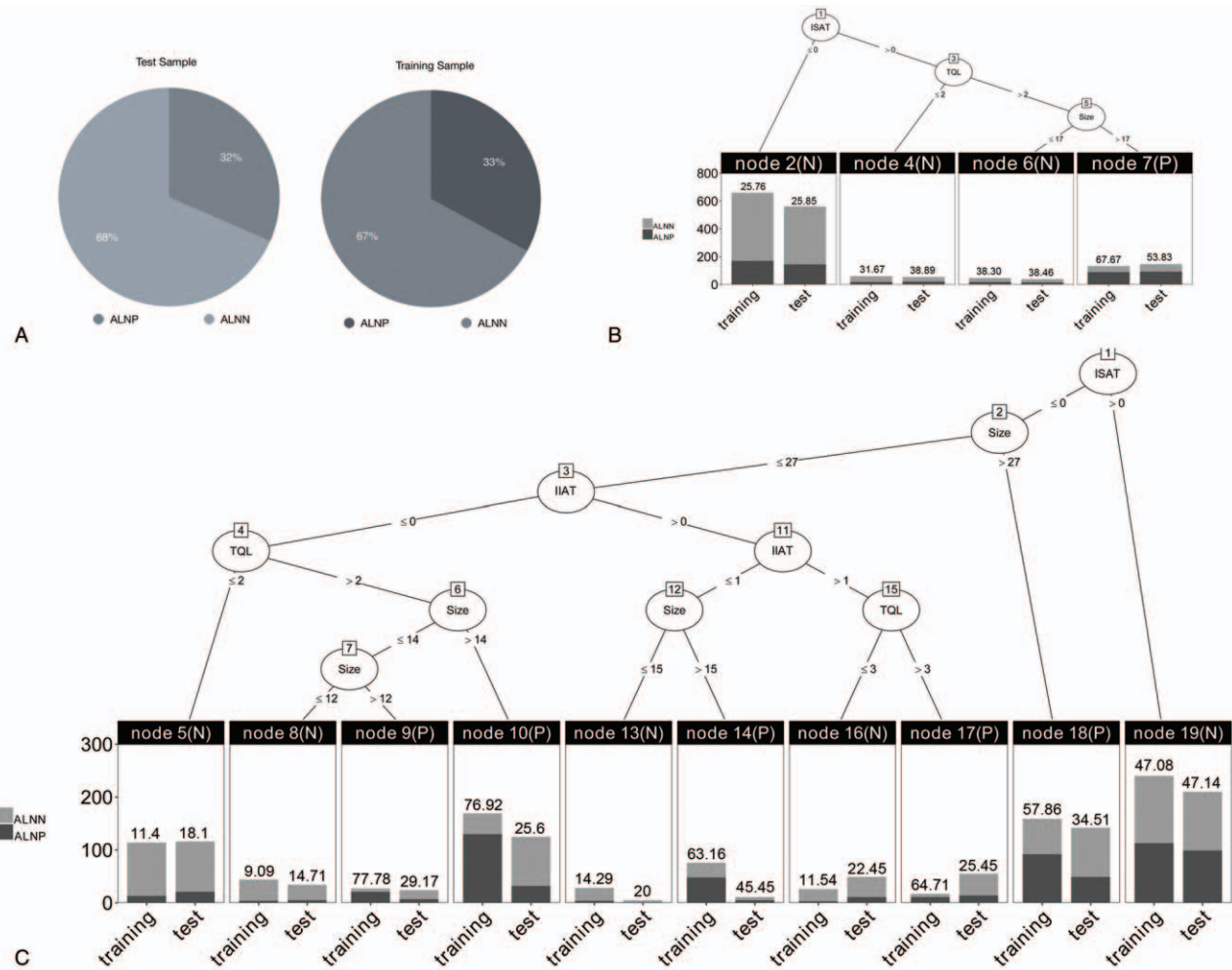


Figure 3. Predictive decision tree models. a. Predictive effectiveness of two data sets. b. Prediction of each node in the accuracy model. The branches above the histogram show the grouping process, and the black cells in the middle represent the prediction results made by the decision tree for each node (N=negative, P=positive). The bottom histogram shows the ALNS distribution of each node; the number on the bar chart represents the percentage of ALNM of each node. c. Prediction of 10 nodes in the error cost model. The detailed explanation is the same as b.

TQLs (26.67%), and tumor size (20.00%), and the model had high accuracy (training sample, 72.2%; test sample, 69.52%; Table 3). Although the accuracy is relatively high in the current model, the FNR was not acceptable (training sample, 69.70%; test sample, 74.18%; Fig. 3b) because the aim of the decision tree was to screen out patients with a risk of ALNM as much as possible so that almost all remaining patients had non-ALNM and might avoid axillary management. Cost matrix can set the severity of each error relative to any other errors; thus, we introduced the cost matrix algorithm setting to a false-positive rate (FPR) 4 times of the false-negative rate (FNR) to reduce FNR.^[27] Then, an error-cost model was established using ISAT (100.00%), tumor size (73.33%), IIAT (55.67%), and TQLs (44.11%). Although the accuracy is lower than that in the accuracy model, the FNR significantly decreased (training sample, 8.08%; test sample, 14.75%; Table 3). Particularly, nodes 5, 8, 13, and 16 predicted non-ALNM with low FNR (training sample, 11.4%, 9.09%, 14.28%, and 11.53%, respectively; test sample, 18.1%, 14.7%, 20%, and 22.45%, respectively, X).

4. Discussion

The timing and distribution of breast cancer metastasis vary considerably. Primary breast tumor metastasis can occur at an early, pre-symptomatic stage, and ALNM is commonly the earliest clinical presentation.^[1] Several studies have confirmed that ALNM in early stage breast cancer can be treated via radiotherapy instead of extended SLNB/ALND.^[2,3] Therefore, exploring and developing non-invasive methods to assess ALNS that will be helpful to avoid unnecessary axillary surgery is a key goal in breast cancer research. In our study, we analyzed several preoperative clinicopathological characteristics and ultrasonographic traits that have been previously reported to be related to ALNM in early stage invasive carcinomas.^[13–19] Finally, we found that tumor size, ISAT, IIAT, and TQLs were significantly correlated with ALNM.

ISAT was the most influential factor affecting ALNM in both multivariate logistic regression analysis and decision tree models. Tumors with ISAT were 2.72 times more likely to develop ALNM than those without ISAT. Lymphatic capillaries have been shown to regulate lymph fluid absorption. Previous studies have

Table 3
Cross-table of predicted and actual axillary lymph node status in the decision models.

	Training sample			Test sample		
	Actual N	Actual P	Actual total	Actual N	Actual P	Actual total
Accuracy model						
Predicted N	TN 560	FN 207	767	TN 473	FN 181	654
Predicted P	FP 43	TP 90	133	FP 54	TP 63	117
Predicted total	603	297	900	527	244	771
False negative rate		69.70%			74.18%	
Accuracy rate		72.22%			69.52%	
Error-cost model						
Predicted N	TN 188	FN 24	212	TN 154	FN 36	190
Predicted P	FP 415	TP 273	688	FP 373	TP 208	581
Predicted total	603	297	900	527	244	771
False negative rate		8.08%			14.75%	
Accuracy rate		51.22%			46.95%	

FN = false negative, FP = false positive, N = axillary lymph node negative, P = axillary lymph node positive, TN = true negative, TP = true positive.

False negative rate (FNR) = FN/(TP+FN)

Accuracy rate = (TP + TN)/(TP + FP + TN + FN)

confirmed that lymphatic capillaries (also known as initial lymphatic capillaries) have no valves, a basement membrane, a smooth muscle cell, or pericyte coating, and its distribution under the skin and subcutaneous adipose tissue is markedly higher than that in the gland and posterior interstitial structure.^[28,29] Thus, preoperative evaluation of ALNS should include assessment of ISAT or NISAT on US. Once ISAT is discovered, clinicians should be vigilant and perform further tests or SLNB to determine ALNS.

Previous studies have shown that tumor size is a predictor of ALNS,^[13–16] and it can reflect the tumor's proliferative capacity to some extent. Studies have demonstrated that bigger tumors are more aggressive and have poorer prognosis.^[13,16] Accordingly, this factor serves as the basis of major staging systems.^[14] In our study, an increase in tumor size by 1 mm increases the risk of ALNM by 1.048, and a larger tumor size was strongly correlated with ALNM. Currently, the T stage is represented by the largest diameter of the invasive cancer component and is easy to obtain. However, it may not be representative of the real tumor burden, particularly in early stage patients. In the future, multidimensional parameters should be considered to obtain a more objective stage based on tumor size.^[30]

TQLs was also suggested to be an influencing factor of ALNM,^[31,32] but no consensus has been achieved to date. Turan et al^[32] explored the clinicopathological factors affecting sentinel lymph node metastasis and found that only lymphovascular invasion could be a useful marker in predicting sentinel lymph node metastasis in patients with axilla-negative early stage breast cancer, while other variables including TQLs have no effect on sentinel lymph node metastasis. Nevertheless, Manjer et al^[31] reported that compared with inner tumors, outer tumors were more strongly associated with a significant risk of ALNM, and central tumors were also significantly associated with ALNM. These findings are consistent with those in our current research. Breast lymph drainage has 2 main routes, namely, the external and internal route. The external route comprises the nipple, the integuments, and the lactiferous tubules and leads to the axilla. Meanwhile, the internal route comprises the dorsal part of the

breast and is thought to perforate the pectoral and intercostal muscles.^[32] Thus, outer tumors and tumors beneath the nipple may reach the axilla through the external route. For IIAT, they may spread through the internal route and reach the lymph nodes behind the pectoralis muscle or intercostal muscles.

Although these factors were connected with ALNS in our study, they cannot be used in clinical practice alone. To determine whether ALNM can be predicted by combining these four indicators, we developed a decision tree. In our accuracy model, the decision tree was built using the ISAT, TQLs, and tumor size and showed a relatively high accuracy rate and high FNR. The FNR of ALNS is a crucial issue, and we attempted to decrease the FNR to the maximum extent possible in clinical practice. After using cost matrix algorithm, the FNR decreased to 14.75% in the test sample.

SLNB as a standard method to assess ALNS can significantly reduce axillary damage to patients with clinical node-negative breast cancer and has a low FNR of approximately 10%.^[33] However, SLNB remains risky for non-ALNM patients and incurs higher healthcare costs. Moreover, the appropriateness of ALND/SLNB for small breast tumor patients with ALNM is yet to be determined,^[3,34] but the results of the AMAROS trial and the ongoing SOUND trial are expected to significantly clarify the role and importance of noninvasive evaluation of ALNS. In our study, we especially focused on the predictive power of the model in non-ALNM patients, for whom surgical staging of the axilla may be eliminated. Although the current model cannot achieve the same FNR compared with SLNB, the FNR was lower than 20% in nodes 5, 8, and 13. This means that these patients might have a higher probability of non-ALNM and may derive no benefit from SLNB or ALND. We speculated that node 5 patients may be at higher risk of internal mammary lymph node metastasis as node 5 tumors are located in the medial quadrant, which is rich in internal mammary lymphatic vessels, and are thus more inclined to drain to the intramammary lymph.^[34] However, our study was focused only on ALNS, and this was not investigated. While for nodes 9, 10, 14, and 18 predicted ALNM, FPR was high and the accuracy significantly decreased because of

the “error-cost algorithm”. So we recommend routinely invasive assessment (SLNB or ALND) for these patients.

There are several limitations of our study. First, it is a retrospective review that only enrolled patients from a single center; thus, the generalizability of the findings is limited. Second, the sample size for the decision tree is relatively small. However, to the best of our knowledge, our study is the first to include LB in a predictive model for evaluating ALNS. In addition, this is the first study to develop an accuracy decision tree model and an error-cost model to predict ALNS in clinical node-negative early stage breast cancer. Although the models are not superior to SLNB, they can provide evidence on the probability of noninvasive measures to evaluate the ALNS and provide a basis for future studies. Basing on the decision tree, clinicians may reference the prediction results and make better clinical decisions.

5. Conclusion

In cT1-T2N0M0 breast cancer, tumors with larger size, lateral TQLs, ISAT, and IIAT may have a higher risk of ALNM. The decision tree built using these independent factors reached a slightly higher FNR than SLNB for predicting ALNS in some patient subgroups. Further advances in diagnostic techniques and systemic therapies for breast cancer are likely to be achieved, and the need for surgical staging of the axilla in select patient subgroups with node-negative disease and some patients with node-positive disease may be eliminated.

Acknowledgments

The authors thank X.H. Mao for his assistance in the statistical analysis.

Author contributions

Conceptualization: Yuanxin Zhang, Xiaomin Li.

Data curation: Yuan Fan.

Funding acquisition: Xiaomin Li.

Investigation: Yuanxin Zhang.

Methodology: Juanjuan Qiu, Mou Zhu.

Project administration: Hongjiang Li.

Resources: Yuanxin Zhang, Xiaomin Li.

Software: Li Ji, Juanjuan Qiu.

Supervision: Hongjiang Li.

Validation: Yuan Fan.

Visualization: Juanjuan Qiu, Mou Zhu.

Writing – original draft: Yuanxin Zhang.

Writing – review & editing: Yuanxin Zhang, Li Ji.

References

- [1] Schwartz RS, Erban JK. Timing of metastasis in breast cancer. *N Engl J Med* 2017;376:2486.
- [2] Galimberti Viviana , Cole Bernard F, Viale Giuseppe , 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.
- [3] Donker Mila , van Tienhoven Geertjan , Straver Marieke E, et al. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): a randomised, multicentre, open-label, phase 3 non-inferiority trial. *Lancet Oncol* 2014;15:1303–10.
- [4] Leone BA, Leone J, Leone JP. Breast cancer is a systemic disease rather than an anatomical process. *Breast Cancer Res Treat* 2017;161:619.
- [5] Redig AJ, McAllister SS. Breast cancer as a systemic disease: a view of metastasis. *J Intern Med* 2013;274:113–26.
- [6] Williams PA, Suggs J, Mangana SH. Axillary lymph node treatment in breast cancer: an update. *J Miss State Med Assoc* 2014;55:145–7.
- [7] Rao R, Euhus D, Mayo HG, et al. Axillary node interventions in breast cancer: a systematic review. *JAMA* 2013;310:1385–94.
- [8] Veronesi U, Paganelli G, Viale G, et al. A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer. *N Engl J Med* 2003;349:546–53.
- [9] Krag DN, Anderson SJ, Julian TB, et al. Primary outcome results of NSABP B-32, a randomized phase III clinical trial to compare sentinel node resection (SNR) to conventional axillary dissection (AD) in clinically node-negative breast cancer patients. *J Clin Oncol* 2010;28 (18_suppl):LBA505.
- [10] Goyal A, Newcombe RG, Mansel RE. Role of routine preoperative lymphoscintigraphy in sentinel node biopsy for breast cancer. *Eur J Cancer* 2005;41:0–243.
- [11] Clarke D, Khonji NI, Mansel RE. Sentinel node biopsy in breast cancer: ALMANAC trial. *World J Surg* 2001;25:819–22.
- [12] Gentilini O, Veronesi U. Abandoning sentinel lymph node biopsy in early breast cancer? A new trial in progress at the European Institute of Oncology of Milan (SOUND: sentinel node vs observation after axillary Ultrasound). *Breast* 2012;21:678–81.
- [13] Yun SJ, Sohn YM, Seo M. Risk stratification for axillary lymph node metastases in breast cancer patients: what clinicopathological and radiological factors of primary breast cancer can predict preoperatively axillary lymph node metastases? *Ultrasound Quart* 2017;33:15–22.
- [14] Kasangian AA, Gherardi G, Biagioli E, et al. The prognostic role of tumor size in early breast cancer in the era of molecular biology. *PLoS One* 2017;12:e0189127.
- [15] Bevilacqua J, Rd CH, Macdonald KA, et al. A prospective validated model for predicting axillary node metastases based on 2,000 sentinel node procedures: the role of tumour location [corrected]. *Eur J Surg Oncol* 2002;28:490–500.
- [16] Bandera EV, John EM. Obesity, body composition, and breast cancer: an evolving science. *JAMA Oncol* 2018;4:804–5.
- [17] Hu X, Zhou X, Yang H, et al. Axillary ultrasound and fine needle aspiration biopsy in the preoperative diagnosis of axillary metastases in early-stage breast cancer. *Oncol Lett* 2018;15:8477–83.
- [18] Evans A, Rauchhaus P, Whelehan P, et al. Does shear wave ultrasound independently predict axillary lymph node metastasis in women with invasive breast cancer? *Breast Cancer Res Treat* 2014; 143:153–7.
- [19] Tate JJ, Lewis V, Archer T, et al. Ultrasound detection of axillary lymph node metastases in breast cancer. *Eur J Surg Oncol* 1989;15:139–41.
- [20] Bae MS, Shin SU, Song SE, et al. Association between US features of primary tumor and axillary lymph node metastasis in patients with clinical T1-T2N0 breast cancer. *Acta Radiologica* 2017; 028418511772303.
- [21] Shetty MK. Letter from the guest editor: breast cancer screening and diagnosis: state of the art. *Semin Ultrasound CT MR* 2011;32:265.
- [22] Gundry KR. Breast ultrasound: indications and findings. *Clin Obstet Gynecol* 2016;59:380.
- [23] Holzer-Fruehwald L, Meissnitzer M, Weber M, et al. Can cut-off-values for tumor size or patient age in breast ultrasound reduce unnecessary biopsies or is it all about bi-rads?—a retrospective analysis of 763 biopsied T1-sized lesions. *Ultrasound Int Open* 2017;03:E94–8.
- [24] Moret BME. Decision Trees and Diagrams. *ACM Computing Surveys* 1982;14:593–623.
- [25] Dietterich TG. Learning at the knowledge level. *Mach Learn* 1986; 1:287–315.
- [26] Hothorn T, Zeileis HA. Unbiased recursive partitioning: a conditional inference framework. *J Comput Graph Stat* 2006;15:651–74.
- [27] Amini P, Ahmadiania H, Poorolajal J, et al. Evaluating the high risk groups for suicide: a comparison of logistic regression, support vector machine, decision tree and artificial neural network. *Iran J Public Health* 2016;45:1179–87.
- [28] Shayan R, Inder R, Karnezis T, et al. Tumor location and nature of lymphatic vessels are key determinants of cancer metastasis. *Clin Exp Metast* 2013;30:345–56.
- [29] Yu H, Zhang S, Zhang R, et al. The role of VEGF-C/D and Flt-4 in the lymphatic metastasis of early-stage invasive cervical carcinoma. *J Exp Clin Cancer Res* 2009;28:98.
- [30] Yu HJ, Chen JH, Mehta RS, et al. MRI measurements of tumor size and pharmacokinetic parameters as early predictors of response in breast

- cancer patients undergoing neoadjuvant anthracycline chemotherapy. *J Magn Reson Imaging* 2010;26:615–23.
- [31] Manjer J, Balldin G, Garne JP. Tumour location and axillary lymph node involvement in breast cancer: a series of 3472 cases from Sweden. *Eur J Surg Oncol* 2004;30:0–617.
- [32] Tanis PJ, Nieweg OE, Valdés Olmos RA, et al. Anatomy and physiology of lymphatic drainage of the breast from the perspective of sentinel node biopsy. *J Am Coll Surg* 2001;192:399–409.
- [33] Krag DN, Anderson SJ, Julian TB, et al. Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial. *Lancet Oncol* 2007;8:881–8.
- [34] Sávolt A, Mátrai Z, Polgár CS, et al. 76. Optimal treatment of the axilla after positive sentinel lymph node biopsy in primary invasive breast cancer patients (surgery versus radiotherapy)–OTOASOR trial: 5 years follow-up of a randomized clinical trial. *Eur J Surg Oncol* 2014;40:S37–8.