



Preoperative prediction of axillary lymph node metastasis in patients with breast cancer based on radiomics of gray-scale ultrasonography

Wei-Jun Zhou^{1,2}, Yi-Dan Zhang³, Wen-Tao Kong³, Chao-Xue Zhang², Bing Zhang^{1,4}

¹Department of Radiology, Nanjing Drum Tower Hospital Clinical College of Nanjing Medical University, Nanjing, China; ²Department of Ultrasound, The First Affiliated Hospital of Anhui Medical University, Hefei, China; ³Department of Ultrasound, Nanjing Drum Tower Hospital Clinical College of Nanjing Medical University, Nanjing, China; ⁴Institute of Brain Science, Nanjing University, Nanjing, China

Contributions: (I) Conception and design: B Zhang; (II) Administrative support: CX Zhang; (III) Provision of study materials or patients: WT Kong; (IV) Collection and assembly of data: WJ Zhou, YD Zhang; (V) Data analysis and interpretation: WJ Zhou; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Bing Zhang. Department of Radiology, Nanjing Drum Tower Hospital Clinical College of Nanjing Medical University, 321, Zhongshan Rd, Nanjing 210008, China. Email: zhangbing_nanjing@nju.edu.cn.

Background: To investigate the performance of a radiomics model based on gray-scale ultrasonography (US) for the preoperative non-invasive prediction of ipsilateral axillary lymph node (ALN) metastasis in patients with breast cancer (BC).

Methods: A total of 192 pathologically confirmed BC patients were included in this study. The training set was comprised of 132 patients from hospital 1 and the test set was comprised of 60 patients from hospital 2. All patients underwent US before percutaneous core biopsy and the results of ALN status reported by a radiologist with 12 years of experience were recorded. Radiomic features were extracted from the gray-scale US images. Max-relevance and min-redundancy (MRMR) and least absolute shrinkage and selection operator (LASSO) were used for data dimension reduction and feature selection. A radiomics model was constructed using LASSO and was validated using the leave group out cross-validation (LGOCV) method. The performance of the model was validated with receiver operating characteristic (ROC), calibration curve, and decision curve analysis.

Results: A total of 860 features were extracted from the gray-scale US images of each breast lesion, and 9 radiomic features were selected for model construction. The area under the curve (AUC), sensitivity, and specificity of the model for predicting ALN metastasis were 0.85, 78.9%, and 77.3% in the training set and 0.65, 68.0%, and 79.4% in the test set, respectively. The prediction performance of the model was significantly higher than that of the radiologist (AUC: 0.85 *vs.* 0.59, $P < 0.01$) in the training set and was slightly higher than that of the radiologist (AUC: 0.65 *vs.* 0.63, $P > 0.05$) in the test set.

Conclusions: The non-invasive radiomics model has the ability to predict ALN metastasis for patients with breast cancer and may outperform US-reported ALN status performed by the radiologist.

Keywords: Radiomics; axillary lymph node (ALN); breast cancer (BC); ultrasound; gray-scale

Submitted Mar 31, 2021. Accepted for publication Jun 04, 2021.

doi: 10.21037/gs-21-315

View this article at: <https://dx.doi.org/10.21037/gs-21-315>

Introduction

Patients with breast cancer (BC) who have axillary lymph node (ALN) involvement, especially involving more than

three nodes, have poorer outcomes than those without node involvement (1,2). There were 30.2–69.8% BC patients have lymph node metastases (3). Therefore, the

recognition of ALN metastasis is crucial for clinical decision making and determining the prognosis of patients with BC. Historically, ALN dissection (ALND) was considered a reliable technique for assessing axillary lymph nodal involvement (4) but because of side effects such as lymphedema, pain, limitation of shoulder movement, and weakness of the arm, ALND was replaced by sentinel lymph node biopsy (SLNB) (2). However, clinical practice has shown that a significant number of BC patients undergo a secondary ALND when SLNB reveals major lymph node involvement (5), leaving no doubt that the accurate preoperative noninvasive assessment of ALN status is essential.

Preoperative noninvasive ALN assessment methods include axillary ultrasonography (US), magnetic resonance imaging (MRI), and mammography. Axillary US is the best method for assessment of preoperative noninvasive ALN (6) and is the safest and most feasible option for ALN staging (7). Axillary US can evaluate nodal morphology real-time, guide fine-needle biopsies, and was useful for predicting pN2 or higher ALN disease in BC patients (8). While traditional US mainly obtains visual image information and focuses on the qualitative analysis of tumors (2), the diagnostic performances of US to detect ALN involvement varies between studies (8-11). Given the limited ability of traditional US to evaluate ALN involvement, developing an applicable non-invasive method for detecting ALN metastasis has been a challenge.

Radiomics, which first proposed by Lambin *et al.* (12), developed rapidly based on medical big data and artificial intelligence and became hotspot of research. It can obtain high-throughput data from biomedical images, extract and analyze quantitative image features, and help quantify indistinguishable differences between tissues in biomedical images (12). In general, the workflow of radiomics include tumor segmentation, feature extraction, feature selection, model construction and model validation. Radiomic features contain first-order, shape, texture and wavelet feature. Radiomics may potentially improve diagnostic, prognostic, and predictive accuracy. Previous studies (13,14) have shown that radiomic feature in primary breast lesions and peritumoral tissue are closely related to tumor heterogeneity. To our knowledge, researches have mainly focused on MRI texture feature analysis (15,16) and there were few reports on US radiomics analysis. Sim *et al.* (17) developed a radiomics-based classifier to distinguish phyllodes tumor and fibroadenoma on gray-scale US images and achieved an area under the

receiver operating characteristic curve (AUC) of 0.765 in the validation set. This research showed the radiomic features extracted from US images are also relate to tumor heterogeneity. But there was no external validation in this study and the generalization capability of the model was not clear. To investigate the relationship between radiomic features based on US images of primary breast tumors and ipsilateral ALN metastasis, we developed a radiomics model based on gray-scale US images for the prediction of ALN metastasis in BC patients. Moreover, we carried out external validation for the model.

We present the following article in accordance with the TRIPOD reporting checklist (available at <https://dx.doi.org/10.21037/gs-21-315>).

Methods

All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional ethics board of the Nanjing Drum Tower Hospital (hospital 1) and the First Affiliated Hospital of Anhui Medical University (hospital 2) (No. 20200171). Individual consent for this retrospective analysis was waived.

Participants and clinical data

The training set of this study was comprised of 132 female patients with histologically confirmed BC from hospital 1 between December 2016 and December 2018. A total of 60 female BC patients were selected from hospital 2 as the test set according to the same criteria from January 2019 to December 2019. All patients underwent ultrasonographic evaluation of the breast tumor and axilla ipsilateral to the tumor and the images were stored in the ultrasonic diagnostic instrument in DICOM form. The patient's age, US-reported tumor size, histologic grade, histologic type, postoperative lymph node status, and the status of HER-2, PR, and ER were retrieved.

The inclusion criteria were (I) the lesion had undergone percutaneous core biopsy and was pathologically confirmed BC; (II) US examination of the breast lesion and ALN was performed within 2 weeks of biopsy; (III) the gray-scale image of the lesion in the largest long axis cross section was available.

The exclusion criteria were (I) patients had undergone preoperative therapy (radiotherapy or chemotherapy) before US examination; (II) the target lesion was not completely

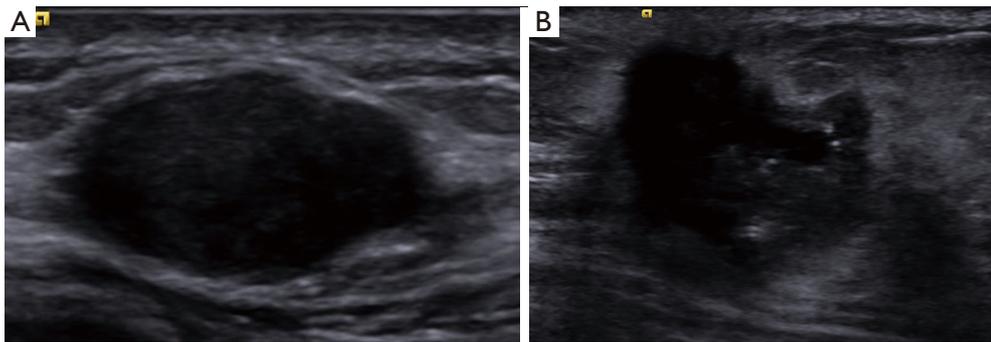


Figure 1 The largest long axis cross-section of breast tumor. (A) Mucinous carcinoma (B) invasive ductal carcinoma

visible in the US image.

US examinations

Gray-scale US images were obtained with the ACUSON S3000 system (Siemens Medical Solutions, Mountain View, CA, USA) and 9L4 linear array transducer (frequency range of 4–9 MHz, Siemens Medical Solutions) by a radiologist with 12 years of experience in breast US. The largest long axis cross-section for each breast tumor was routinely recorded (*Figure 1*).

US-reported ALN status was assessed according to the following criteria (18): (I) normal ALN present as oval or reniform in shape, with a thin (thickness is less than 3 mm), even, smooth, C-shaped hypoechoic cortex, and a hyperechoic central fatty hilum; (II) ALN for suspicious metastasis present as round shape, focal or eccentric thickened cortex (thickness is more than 3 mm), peripheral vascularization, and an indented or effaced fatty hilum. The radiologist was blinded to the clinical and pathologic information of patients.

Region of interest (ROI) segmentation

The ROI of the whole primary breast tumor on the gray-scale images of the largest long axis cross-section was manually segmented using open-source software ITK-SNAP 3.6.0 (<https://www.itksnap.org>) by the radiologist with 12 years of experience in breast US. Thirty patients were then randomly selected, and the ROI was delineated by another radiologist with 5 years of experience in breast US to assess the consistency of the readers. The workflow of tumor segmentation, radiomic features extraction, features extraction, model construction, and model validation are shown in *Figure 2*.

Feature extraction and selection

Firstly, we used a Z-score method to normalize each gray-scale image of the breast tumors, then radiomic features were extracted from each image using the MATLAB 2013a platform (MathWorks, Natick, USA). The radiomic features were classified into four categories of features, including first-order, shape, texture, and wavelet features (19). The inter-observer reproducibility of radiomic feature extraction were evaluated with intra-class correlation coefficient (ICC). When the ICC value reached 0.75, further analysis was performed using each feature (20).

Secondly, max-relevance and min-redundancy (MRMR) (21) was performed to reduce the dimension of the radiomics features and eliminate redundant and irrelevant features. Thirdly, least absolute shrinkage and selection operator (LASSO) (22) were used to select the most relevant features and LASSO constrain was used as the classifier to construct the final radiomics model. LASSO chose the optimal penalization coefficient parameter λ via 10-fold cross validation and determined the number of the feature by optimal λ in a coefficient profile plot on the training set (*Figure 3A,B*).

Model construction

The radiomics score (Rad-score) for each breast tumor was calculated by summing the selected features weighted by their coefficients. The Rad-scores of patients with and without ALN metastasis on the training set and test set were compared.

Model validation

In this study, the validity of the prediction model was

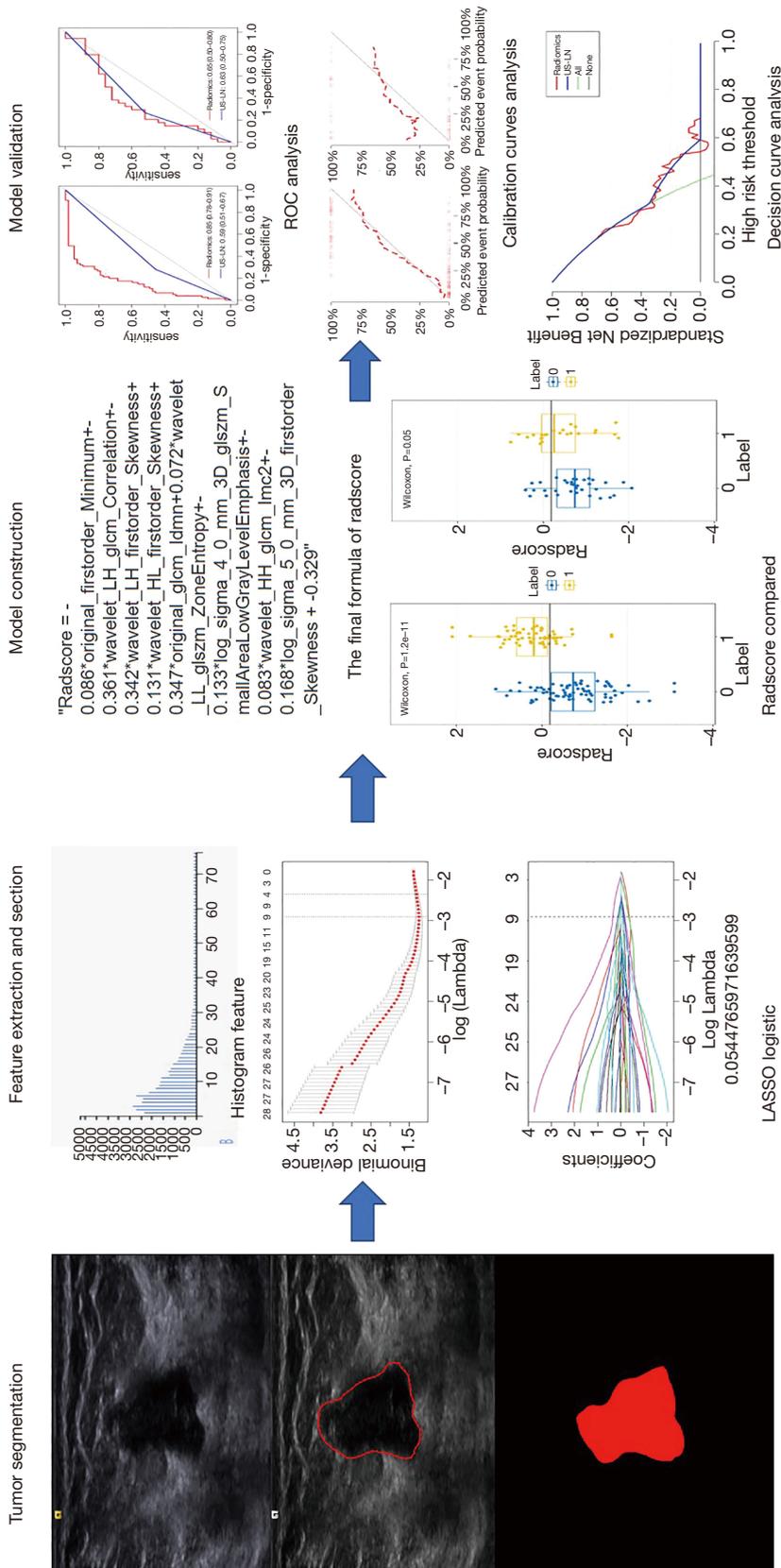


Figure 2 Flowchart of the study. First, the target tumor was manually segmented using ITK-SNAP. Next, a total of 860 features were extracted from each image in our study with MATLAB. MRM and the LASSO were used for feature selection. In LASSO regression, a coefficient profile plot was drawn and resulted in nine radiomic features with nonzero coefficients. Thereafter, Rad-score was calculated by the selected nonzero coefficients features. Finally, the model was validated with ROC) calibration curve and decision curve analysis. MRM, max-relevance and min-redundancy; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic.

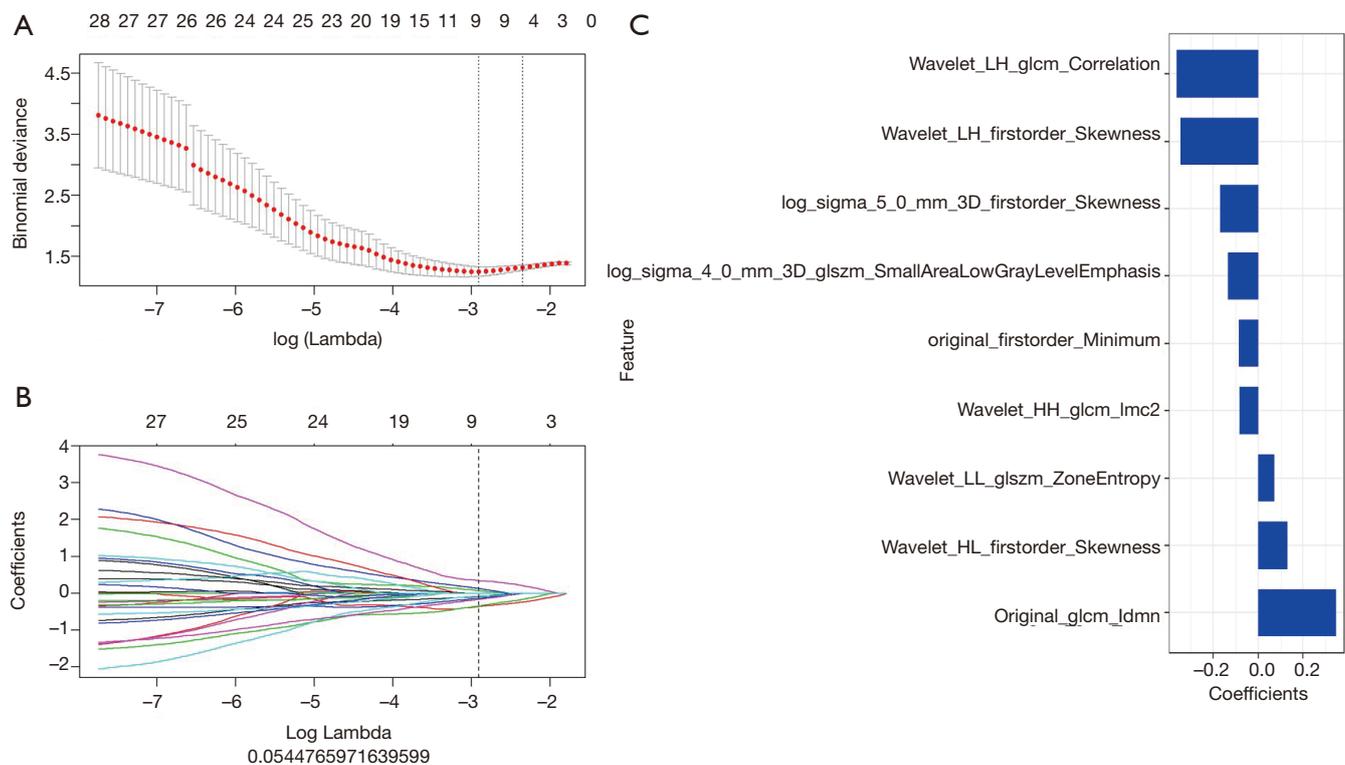


Figure 3 Tuning parameter selection using the LASSO in the training set. (A) The optimal penalization coefficient λ was generated in the LASSO via 10-fold cross validation. The λ value of the minimum mean binomial deviation for the training set was given for the feature with zero selection coefficient. (B) LASSO coefficient profiles of the 860 features. L1 norm was added in the final lost function and the weights were constrained, which makes the features sparse. To determine the hyper-parameter (e.g., the number of features) of the model, we applied cross validation with 10-fold on the training data set. (C) In LASSO regression, a coefficient profile plot was drawn and resulted in 9 radiomic features with nonzero coefficients. LASSO, least absolute shrinkage and selection operator.

verified by discrimination, calibration, and clinical application.

Discrimination

The pathological results of ALNs were taken as the diagnostic criteria. The receiver operating characteristic (ROC) curve was plotted based on the pathological results of ALNs to show the performance for ALN metastasis prediction of the Rad-score and the radiologist in the training and test sets, and the areas under the curve (AUCs) were obtained to quantify the discriminating ability. In addition, we performed 100 times leave group out cross-validation (LGOCV) to verify the reliability of our results. The mean AUCs of training and validation sets were calculated, and the optimal cutoff value of the Rad-score which was calculated in the training set was applied to the test set for verification. The performance between the

model and the radiologist were then compared.

Calibration

The calibration curve was plotted to explore the predictive accuracy of the model in the training and test sets. The calibration curves determined the coincidence between the predicted probability of ALN metastasis and the actual outcomes (i.e., the ALN metastasis) in both training and test sets. To evaluate whether the model was well-calibrated or not, the Hosmer-Lemeshow statistics were calculated to reflect the reliability of the calibration curves, and a P value greater than 0.05 was considered to be a good calibration.

Clinical application

The clinical decision curve analysis (DCA) was plotted to show the clinical effectiveness of the model by quantifying the net benefits at different threshold probabilities in the

training and test sets.

Statistical analysis

Statistical analysis was performed with IBM SPSS 22.0 (IBM Corp, Armonk, New York) and R software (version 4.0.4; <http://www.Rproject.org>). $P \leq 0.05$ (two-sided) was considered to be statistically significant. Continuous variables such as age were represented as mean \pm standard deviation, categorical variables such as histological grade and histologic types were represented as N (%), and the Rad-score was represented as the median (interquartile range).

SPSS 22.0 was used to compare the variables between different cohorts. Continuous variables with a normal distribution such as age and US-reported tumor size were compared using the *t*-test, and continuous variables with an abnormal distribution such as Rad-score were compared using the Mann-Whitney U test. Categorical variables were compared using the χ^2 test.

R software was used for building and evaluating the prediction model. The LASSO analysis, LGOCV analysis, ROC curve, calibration curve, and decision curve analysis were performed using “glmnet”, “caret”, “pROC”, “Calibration Curves”, and “Decision Curve” packages, respectively.

Results

Clinical and pathological characteristics

A total of 192 BC patients, including patients with ALN metastasis ($n=83$) and without ALN metastasis ($n=109$), were included in this study. The incidence of ALN metastasis in BC patients was 43.2%. All patients were divided into a training set ($n=132$, from hospital 1) and test set ($n=60$, from hospital 2). The incidence of ALN metastasis in training and test set was 43.2% (57/132) and 43.3% (26/60), respectively. The clinical and pathological features of the training and test sets were compared (Table 1). There were no significant differences in the frequency of ALN metastasis, US-reported tumor size, histological grade, pathological type, HER-2, ER, PR, histological grade, and US-reported ALN status between the training and test sets (all $P > 0.05$).

We also investigated the above characteristics between patients with and without ALN metastasis in the training and test sets (Table 2). There were statistically significant differences between patients with and without ALN metastasis in histological type, US-reported ALN status,

and Rad-score ($P < 0.05$) in the training set. In the test set, there was statistically significant difference in US-reported ALN status and Rad-score between patients with and without ALN metastasis ($P < 0.05$). The accuracy of the US-reported ALN status in the training and test sets was 64.4% (85/132) and 61.7% (37/60), respectively, and there was no significant difference in US-reported ALN status between the two sets ($P > 0.05$).

Radiomics feature extraction, selection, and model construction

A total of 860 features were extracted from each image in our study. These features included first-order features ($n=18$), shape features ($n=38$), texture features ($n=460$), and wavelet features ($n=344$). After evaluating the reproducibility of feature extraction and using the MRMR method, there were 30 features retained. Finally, in LASSO regression, a coefficient profile plot was drawn and resulted in 9 radiomic features with nonzero coefficients. These features were wavelet_HL_glcm_Correlation, wavelet_LH_firstorder_Skewness, log_sigma_5_0_mm_3D_firstorder_Skewness, log_sigma_4_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis, original_firstorder_Minimum, wavelet_HH_glcm_Imc2, wavelet_LL_glszm_ZoneEntropy, wavelet_HL_firstorder_Skewness, and original_glcm_Idmn (Figure 3C).

The final formula of the Rad-score was: Rad-score = $-0.086 * \text{original_firstorder_Minimum} + 0.361 * \text{wavelet_LH_glcm_Correlation} + 0.342 * \text{wavelet_LH_firstorder_Skewness} + 0.131 * \text{wavelet_HL_firstorder_Skewness} + 0.347 * \text{original_glcm_Idmn} + 0.072 * \text{wavelet_LL_glszm_ZoneEntropy} + 0.133 * \text{log_sigma_4_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis} + 0.083 * \text{wavelet_HH_glcm_Imc2} + 0.168 * \text{log_sigma_5_0_mm_3D_firstorder_Skewness} + -0.329$.

While there was no significant difference between the training and test sets in Rad-score (median, -0.902 vs. -0.667 ; $P = 0.788$) (Table 1), there was statistically significant difference in Rad-score between the BC patients with and without ALN metastasis in the training set (median, -0.439 vs. -1.020 ; $P = 0.007$), which was verified in the test set (median, -0.254 vs. -1.059 ; $P = 0.033$) (Table 2).

Model validation

Discrimination of the model

The model showed a significant higher discrimination

Table 1 The baseline characteristics of the enrolled patient in the training and test sets

Characteristics	Training set (N=132)	Test set (N=60)	P value
Age (year)	54.07±11.12	54.35±11.38	0.872
US-reported tumor size (cm)	2.31±0.78	2.10±0.71	0.073
Histologic grade, n (%)			0.748
I	17 (12.9)	8 (13.3)	
II	84 (63.6)	35 (58.3)	
III	31 (23.5)	17 (28.3)	
Histologic type, n (%)			0.084
IDC	100 (75.8)	52 (86.7)	
Other	32 (24.2)	8 (13.3)	
HER-2, n (%)			0.121
Positive	62 (47.0)	21 (35.0)	
Negative	70 (53.0)	39 (65.0)	
ER, n (%)			0.405
Positive	104 (78.8)	44 (73.3)	
Negative	28 (21.2)	16 (26.7)	
PR, n (%)			0.312
Positive	87 (65.9)	35 (58.3)	
Negative	45 (34.1)	25 (41.7)	
ALN status, n (%)			0.931
Positive	56 (42.4)	26 (43.3)	
Negative	76 (57.6)	34 (56.7)	
US-reported LN status, n (%)			0.716
LN-suspicious metastasis	47 (35.6)	23 (38.3)	
LN-normal	85 (64.4)	37 (61.7)	
Rad-score	-0.902 (-1.52, -0.27)	-0.667 (-1.64, 0.00)	0.788

Rad-score was represented by median (interquartile range) and $P < 0.05$ indicated that the difference was statistically significant. US, ultrasound; IDC, invasive ductal carcinoma; HER2, human epidermal growth factor receptor; ER, estrogen receptor; PR, progesterone receptor.

performance than the radiologist in the training set [AUC, 0.85 (95% CI: 0.78–0.91) *vs.* AUC, 0.59 (95% CI: 0.51–0.67), $P < 0.001$]. The model also presented a slightly higher discrimination performance than the radiologist in the test set [AUC, 0.65 (95% CI: 0.50–0.80) *vs.* AUC, 0.63 (95% CI: 0.50–0.75), $P > 0.05$] (Figure 4A,B, Table 3). The sensitivity and specificity of the model were higher than US-reported ALN status in the training set (78.9%, 77.3% *vs.* 45.6%, 72.0%) and also verified in the test set (68.0%, 79.4% *vs.* 52%, 73.5%).

The LGOCV was performed 100 times to verify the reliability of the results, and in LGOCV analysis, a high pooled mean AUC (0.864 and 0.734 in the training and validation sets, respectively) suggested that the prediction model achieved better predictive efficacy than the radiologist.

Calibration of the prediction model

The calibration curves of the radiomics model in the training set (Figure 4C) revealed that the model showed

Table 2 Baseline characteristics between the patients with ALN metastasis and without ALN metastasis in the training and test sets

Characteristic	Training set			Test set		
	LN metastasis (+) (n=57)	LN metastasis (-) (n=75)	P	LN metastasis (+) (n=26)	LN metastasis (-) (n=34)	P
Age (years)	55.79±10.18	52.76±11.68	0.121	51.81±10.80	56.29±11.60	0.132
US-reported tumor size (cm)	2.45±0.70	2.22±0.84	0.085	2.25±0.74	1.99±0.68	0.165
Histologic grade, n (%)			0.678			0.167
I	6 (10.5)	11 (14.7)		1 (3.8)	7 (20.6)	
II	36 (63.2)	48 (64.0)		17 (65.4)	18 (52.9)	
III	15 (26.3)	16 (21.3)		8 (30.8)	9 (26.5)	
Histologic type, n (%)			0.017			0.059
IDC	49 (86.0)	51 (68.0)		25 (96.2)	27 (79.4)	
Other	8 (14.0)	24 (32.0)		1 (3.8)	7 (20.6)	
HER-2, n (%)			0.433			0.956
Positive	29 (50.9)	33 (44.0)		9 (34.6)	12 (35.3)	
Negative	28 (49.1)	42 (56.0)		17 (65.4)	22 (64.7)	
ER, n (%)			0.412			0.969
Positive	43 (75.4)	61 (81.3)		19 (73.1)	25 (73.5)	
Negative	14 (24.6)	14 (18.7)		7 (26.9)	9 (26.5)	
PR, n (%)			0.186			0.333
Positive	34 (59.6)	53 (70.7)		17 (65.4)	18 (52.9)	
Negative	23 (40.4)	22 (29.3)		9 (34.6)	16 (47.1)	
US-reported LN status, n (%)			0.036			0.031
LN-suspicious metastasis	26 (45.6)	21 (28.0)		14 (53.8)	9 (26.5)	
LN-normal	31 (54.4)	54 (72.0)		12 (46.2)	25 (73.5)	
Rad-score	-0.439 (-1.309, 0.534)	-1.020 (-1.545, -0.443)	0.007	-0.254 (-1.623, 1.165)	-1.059 (-1.754, -0.309)	0.033

Rad-score was represented by median (interquartile range) and $P < 0.05$ indicated that the difference was statistically significant. US, ultrasound; IDC, invasive ductal breast cancer; HER2, human epidermal growth factor receptor; ER, estrogen receptor; PR, progesterone receptor.

an accurate agreement between the prediction of ALN metastasis in BC patients and observation in the training set. The Hosmer-Lemeshow statistics showed $\chi^2=8.507$, $P=0.386$ in the training set and showed $\chi^2=11.225$, $P=0.189$ in the test set, which revealed there is no statistically significant difference between the predicted value and the actual observed value, and the prediction model has good calibration ability in both the training set and the test set. (Figure 4D).

Clinical application of the prediction model

The decision curve for the prediction model was shown in Figure 5. This showed that if the threshold probability exceeds 32%, the predictive model will obtain more benefit than all treatment protocols (assuming all BC patients had ALN metastases) or no treatment (assuming no ALN metastases in all breast cancer patients). Furthermore, if the threshold probability was between 38% and 53%, or >59%, using the radiomics prediction model to predict ALN

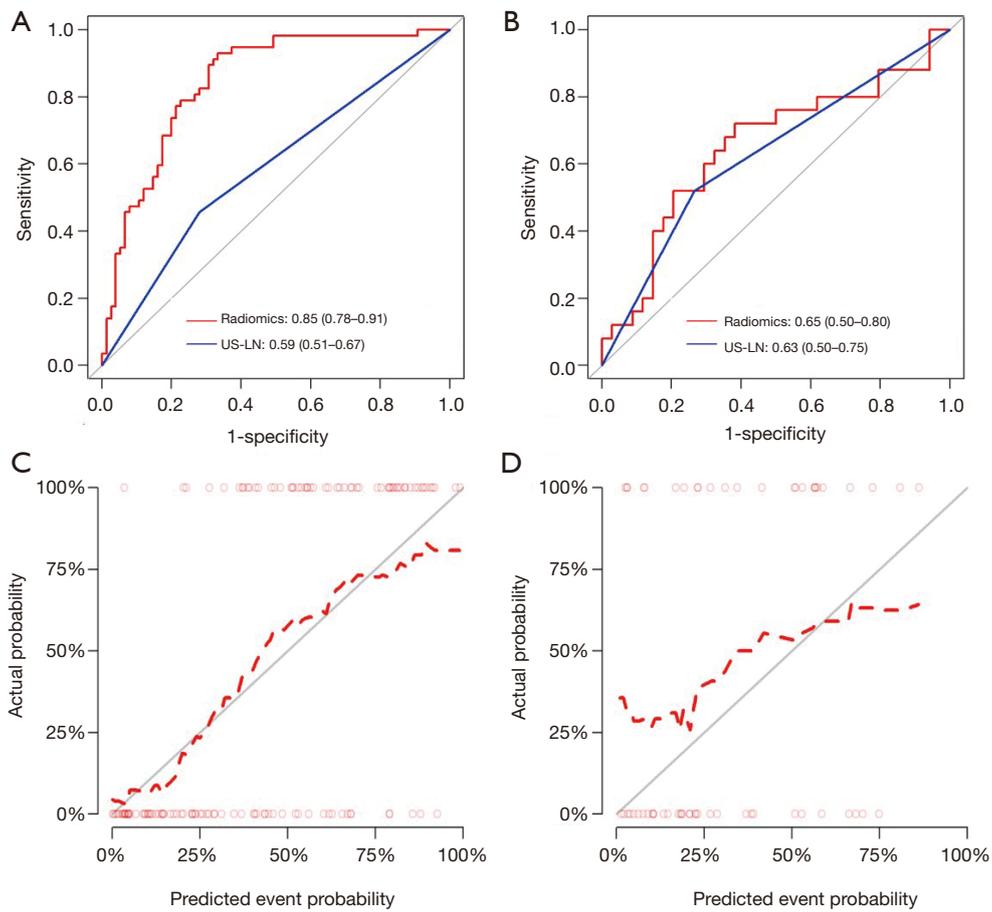


Figure 4 Model validation. (A,B) ROC curves of the radiomics model (red lines) and ALN status reported by a radiologist with 12 years of experience (blue lines) in the training (A) and test sets (B). (C,D) Calibration curves of the radiomics model in the training (C) and test sets (D). ROC, receiver operating characteristic; US, ultrasound; ALN, axillary lymph node.

Table 3 The performance of radiomics model and US-reported ALN status in the training and test sets

Variables	Training set		Test set	
	Radiomics model	US-reported ALN status	Radiomics model	US-reported ALN status
AUC (95% CI)	0.85 (0.78–0.91)	0.59 (0.51–0.67)*	0.65 (0.50–0.80)	0.63 (0.50–0.75)
Threshold	–0.182	NA	–0.182	NA
Sensitivity	78.9%	45.6%	68.0%	52.0%
Specificity	77.3%	72.0%	79.4%	73.5%
Accuracy	78.0%	60.6%	66.1%	64.4%
PPV	72.5%	55.3%	63.1%	59.0%
NPV	82.8%	63.5%	67.5%	67.6%

*, compared with radiomics model in the training set, $P < 0.001$. ALN, axillary lymph node; PPV, positive predictive value; NPV, negative predictive value.

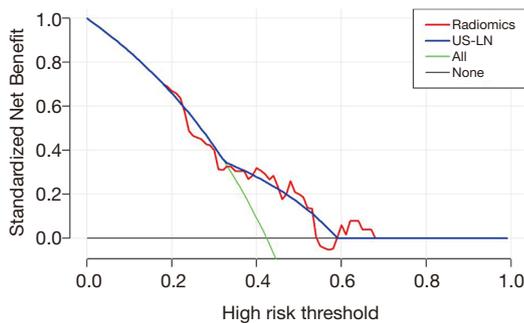


Figure 5 Decision curve of the radiomics predictive model (red line) and the ALN status reported by a radiologist with 12 years of experience (blue line). The y-axis represents the standardized net benefit, the x-axis represents the threshold probability, the black line represents the assumption that no patients showed ALN metastasis, and the green line represents the assumption that all patients showed ALN metastasis. If the threshold probability exceeds 32%, the predictive model will obtain more benefit than all treatment protocols (assuming all breast cancer patients had ALN metastases) or no treatment (assuming no ALN metastases in all breast cancer patients). If the threshold probability was between 38% and 53%, or >59%, using the radiomics prediction model in the study to predict ALN metastases shows a greater advantage than the radiologist. US, ultrasound; ALN, axillary lymph node.

metastases showed a greater advantage than the radiologist.

Discussion

Our study developed a radiomics model based on gray-scale US images of breast cancer for the noninvasive preoperative prediction of ipsilateral ALN metastasis in breast cancer patients. The radiomics model was validated by ROC, calibration, and clinical decision curves in the test set. The AUC, sensitivity and specificity of the model was higher than the radiologist with 12 years of service (0.85, 78.9%, 77.3% vs. 0.59, 45.6%, 72.0%, respectively) in the training set. Furthermore, the results were validated in the test set (AUC, sensitivity, specificity: 0.65, 68.0%, 79.4% vs. 0.63, 52.0%, 73.5%, respectively).

Radiomics can fully reflect the heterogeneity of tumors and has become a reliable potential biomarker for improving diagnosis, treatment decisions, and the prediction of treatment response. Radiomics has usually been performed basing on MRI and computer tomography (CT) images of tumors (13,14,23). Compared with MRI, US has the advantages of convenient operation, real-time

scanning, multi-section continuous dynamic scanning and so on. The combination of US and radiomics may be more favorable for breast cancer screening in china. Moreover, it is encouraging that a growing number of studies have demonstrated that ultrasonic images can also be used for radiomics analysis. In addition, the preoperative evaluation of ALN in breast cancer patients mainly depends on axillary US, Mammography and MRI (24). In generally, mammography is suboptimal for ALN evaluation because of the low sensitivity and accuracy and MRI has the highest sensitivity on ALN assessment (25). Axillary US is widely used to evaluate ALN status. Nonetheless, the results varied according to operator (25). Akissue de Camargo Teixeira *et al.* (26) demonstrated a nomogram model based on ultrasonic images of ALNs can predict ALN metastasis. In our study, radiomics was used to analyze the US images of breast tumors, and 9 features were extracted from each primary breast lesions. Among the 9 features, wavelet_LH_GLCM_Correlation showed the strongest correlation with ALN metastasis, while wavelet_LH_firstorder_Skewness and original_glcmm_Idmn also showed a strong correlation. These radiomics features are higher-order features and represent the heterogeneity of the tumor and the slight difference of the gray and texture features (27). GLCM features have been widely used in many texture analysis applications, and most GLCM parameters could describe the echo pattern homogeneity (28). Some previous studies (29,30) showed GLCM features extracted from T1-weighted images of dynamic contrast enhanced and US images are correlated with LN involvements. In our study, GLCM features extracted from US image of primary breast lesions are also correlated with ALN metastasis. Furthermore, skewness was also shown to be associated with ALN metastasis in our study, which was consistent with Han *et al.* (30), who found skewness could predict ALN metastasis effectively. Tumor features at the genetic and cellular levels can be extracted on biomedical images (12,31). Therefore, radiomics features extracted from US image of primary breast lesions can well distinguish the invisible heterogeneity of tumors and are available to predict the status of ALN.

To improve the diagnostic accuracy, we developed a radiomics model to assess the ALN status and compared the performance of model with a senior radiologist with 12 years' experience in US. The result revealed that the performance of the model was superior to the radiologist both in the training set and in the test set. Lee *et al.* (32) developed a radiomics model base on US features of

primary breast cancers and the performance of the model was superior to the clinicopathologic model. But there was no external validation in this research for the model. Our results were consistent with this study and we did external validation for the radiomics model. The results indicate that represented radiomics model based on US imaging provides a promising tool for predicting axillary lymph node metastasis in patients with breast cancer.

Traditional radiomics analysis mostly uses one feature selection method. Aiming to reduce overfitting effectively, our study adopted MRMR for feature selection in the first step and LASSO for feature selection in the second step. A previous study (33) selected features with LASSO only and developed a radiomics nomogram based on ultrasonic images in primary breast lesions and obtained moderate prediction performance both in primary and validation cohorts. To obtain a better predictive performance, we improved the feature selection method in our study, and finally achieved excellent prediction efficacy with an AUC of 0.85 in the training set and moderate prediction performance with an AUC of 0.65 in the test set. In LGOCV analysis, a high pooled mean AUC (0.864 and 0.734 in training and validation sets, respectively) indicated this is a relatively reliable and stable predictive model. In turn, this indicated that the radiomics feature analysis of primary breast tumors can effectively predict the status of ALNs.

There are several limitations to our study. First, this research is based on gray-scale US images which cannot completely overcome operator dependency and might result in sampling bias. Second, all the US data were obtained from the same type ultrasonic system, and the predicational model was not validated on samples from other instruments. Therefore, the prediction model has not been applied in clinical and the applicability of this model requires further evaluation. Thirdly, our study only analyzed the relationship between the radiomics features of primary breast tumor and the status of ALNs and did not further analyze the radiomics features related to the number of metastatic SLNs. Therefore, in future studies we will further analyze the relationship between the radiomics features of BC and the number of metastatic SLNs. These findings may improve the ability of clinicians to influence therapeutic regimens and clinical decision-making in the management of BC patients.

In conclusion, we developed and validated a novel non-invasive preoperative radiomics model for ALN metastasis based on gray-scale US images of breast primary tumors.

The model has the ability to predict ALN metastasis for patients with breast cancer and may outperform US-reported ALN status performed by radiologists.

Acknowledgments

Funding: This work was supported by the Nanjing Science and Technology commission [grant number 201803027]; the National Natural Science Foundation of China (81720108022 B.Z., 81971596, X.Z., 81701672, Z.Q.); the Social Development Project of Science and Technology Project in Jiangsu Province (BE2017707); Key Medical Talents of the Jiangsu Province, the “13th Five-Year” Health Promotion Project of the Jiangsu Province (ZDRCA2016064); Jiangsu Provincial Key Medical Discipline (Laboratory) (ZDXKA2016020); the Project of the Sixth Peak of Talented People (WSN-138). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript..

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://dx.doi.org/10.21037/gs-21-315>

Data Sharing Statement: Available at <https://dx.doi.org/10.21037/gs-21-315>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/gs-21-315>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional ethics board of the Nanjing Drum Tower Hospital and the First Affiliated Hospital of Anhui Medical University (No.: 20200171). Individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons

Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Tjan-Heijnen V, Viale G. The Lymph Node and the Metastasis. *N Engl J Med* 2018;378:2045-6.
2. Maxwell F, de Margerie Mellon C, Bricout M, et al. Diagnostic strategy for the assessment of axillary lymph node status in breast cancer. *Diagn Interv Imaging* 2015;96:1089-101.
3. Niinikoski L, Hukkinen K, Leidenius MHK, et al. Axillary nodal metastatic burden in patients with breast cancer with clinically positive axillary nodes. *Br J Surg* 2020;107:1615-1624.
4. Giuliano AE, Ballman KV, McCall L, et al. Effect of Axillary Dissection vs No Axillary Dissection on 10-Year Overall Survival Among Women With Invasive Breast Cancer and Sentinel Node Metastasis: The ACOSOG Z0011 (Alliance) Randomized Clinical Trial. *Jama* 2017;318:918-26.
5. Maxwell F, de Margerie Mellon C, Bricout M, et al. Diagnostic strategy for the assessment of axillary lymph node status in breast cancer. *Diagn Interv Imaging* 2015;96:1089-101.
6. Choi HY, Park M, Seo M, et al. Preoperative Axillary Lymph Node Evaluation in Breast Cancer: Current Issues and Literature Review. *Ultrasound quarterly* 2017;33:6-14.
7. Lee MC, Joh JE, Chau A. Axillary staging prior to neoadjuvant chemotherapy: the roles of sentinel lymph node biopsy and axillary ultrasonography. *Cancer Control* 2012;19:277-85.
8. Abe H, Schacht D, Sennett CA, et al. Utility of preoperative ultrasound for predicting pN2 or higher stage axillary lymph node involvement in patients with newly diagnosed breast cancer. *AJR Am J Roentgenol* 2013;200:696-702.
9. Ferreira da Silva F, de Lurdes Orvalho M, Gaspar A, et al. Preoperative axillary lymph node staging in early breast cancer: Diagnostic efficacy and histopathology correlation. *Breast J* 2019;25:1312-5.
10. van Nijnatten TJA, Ploumen EH, Schipper RJ, et al. Routine use of standard breast MRI compared to axillary ultrasound for differentiating between no, limited and advanced axillary nodal disease in newly diagnosed breast cancer patients. *Eur J Radiol* 2016;85:2288-94.
11. Nakamura R, Yamamoto N, Miyaki T, et al. Impact of sentinel lymph node biopsy by ultrasound-guided core needle biopsy for patients with suspicious node positive breast cancer. *Breast Cancer* 2018;25:86-93.
12. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441-6.
13. Wang T, Gao T, Yang J, et al. Preoperative prediction of pelvic lymph nodes metastasis in early-stage cervical cancer using radiomics nomogram developed based on T2-weighted MRI and diffusion-weighted imaging. *Eur J Radiol* 2019;114:128-35.
14. Dong Y, Feng Q, Yang W, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. *Eur Radiol* 2018;28:582-91.
15. Valdora F, Houssami N, Rossi F, et al. Rapid review: radiomics and breast cancer. *Breast Cancer Res Treat* 2018;169:217-29.
16. Chitalia RD, Kontos D. Role of texture analysis in breast MRI as a cancer biomarker: A review. *J Magn Reson Imaging* 2019;49:927-38.
17. Sim Y, Lee SE, Kim EK, et al. A Radiomics Approach for the Classification of Fibroepithelial Lesions on Breast Ultrasonography. *Ultrasound Med Biol* 2020;46:1133-41.
18. Sun SX, Moseley TW, Kuerer HM, et al. Imaging-Based Approach to Axillary Lymph Node Staging and Sentinel Lymph Node Biopsy in Patients With Breast Cancer. *AJR Am J Roentgenol* 2020;214:249-58.
19. Liang C, Cheng Z, Huang Y, et al. An MRI-based Radiomics Classifier for Preoperative Prediction of Ki-67 Status in Breast Cancer. *Acad Radiol* 2018;25:1111-7.
20. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016;15:155-63.
21. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 2005;27:1226-38.
22. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* 2007;26:5512-28.

23. Liu Z, Zhang XY, Shi YJ, et al. Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Clin Cancer Res* 2017;23:7253-62.
 24. Chang JM, Leung JWT, Moy L, et al. Axillary nodal evaluation in breast cancer: state of the art. *Radiology* 2020;295:500-15.
 25. Choi HY, Park M, Seo M, et al. Preoperative Axillary Lymph Node Evaluation in Breast Cancer: Current Issues and Literature Review. *Ultrasound Q* 2017;33:6-14.
 26. Akissue de Camargo Teixeira P, Chala LF, Shimizu C, et al. Axillary Lymph Node Sonographic Features and Breast Tumor Characteristics as Predictors of Malignancy: A Nomogram to Predict Risk. *Ultrasound Med Biol* 2017;43:1837-45.
 27. Huynh E, Coroller TP, Narayan V, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiother Oncol* 2016;120:258-66.
 28. Guo Y, Hu Y, Qiao M, et al. Radiomics Analysis on Ultrasound for Prediction of Biologic Behavior in Breast Invasive Ductal Carcinoma. *Clin Breast Cancer* 2018;18:e335-e344.
 29. Theek B, Opacic T, Magnuska Z, et al. Radiomic analysis of contrast-enhanced ultrasound data. *Sci Rep* 2018;8:11359.
 30. Han L, Zhu Y, Liu Z, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol* 2019;29:3820-9.
 31. Yang L, Dong D, Fang M, et al. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *Eur Radiol* 2018;28:2058-67.
 32. Lee SE, Sim Y, Kim S, et al. Predictive performance of ultrasonography-based radiomics for axillary lymph node metastasis in the preoperative evaluation of breast cancer. *Ultrasonography* 2021;40:93-102.
 33. Yu FH, Wang JX, Ye XH, et al. Ultrasound-based radiomics nomogram: A potential biomarker to predict axillary lymph node metastasis in early-stage invasive breast cancer. *Eur J Radiol* 2019;119:108658.
- (English Language Editor: B. Draper)

Cite this article as: Zhou WJ, Zhang YD, Kong WT, Zhang CX, Zhang B. Preoperative prediction of axillary lymph node metastasis in patients with breast cancer based on radiomics of gray-scale ultrasonography. *Gland Surg* 2021;10(6):1989-2001. doi: 10.21037/gs-21-315