



Development of a nomogram-based model combining intra- and peritumoral ultrasound radiomics with clinical features for differentiating benign from malignant in Breast Imaging Reporting and Data System category 3–5 nodules

Lichang Zhong^{1#}, Lin Shi^{1#}, Liang Zhou², Xinpeng Liu³, Liping Gu¹, Wenkun Bai¹

¹Department of Ultrasound in Medicine, Sixth People's Hospital Affiliated to Medical College of Shanghai Jiao Tong University, Shanghai Institute of Ultrasound in Medicine, Shanghai, China; ²Department of Information, Sixth People's Hospital Affiliated to Medical College of Shanghai Jiao Tong University, Shanghai, China; ³Faculty of Chinese Medicine, Macau University of Science and Technology, Macau, China

Contributions: (I) Conception and design: L Zhong, W Bai; (II) Administrative support: W Bai; (III) Provision of study materials or patients: L Zhong, L Shi, L Gu, W Bai; (IV) Collection and assembly of data: L Zhong, L Shi, L Gu; (V) Data analysis and interpretation: L Zhong, L Zhou, X Liu, W Bai; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Wenkun Bai, PhD. Department of Ultrasound in Medicine, Sixth People's Hospital Affiliated to Medical College of Shanghai Jiao Tong University, Shanghai Institute of Ultrasound in Medicine, No. 600, Yishan Road, Shanghai 200233, China. Email: doctor505@hotmail.com.

Background: The differences in benign and malignant breast tumors are not only within the nodules but also involve changes in the surrounding tissues. Radiomics can reveal many details that are not discernible to the naked eye. This study aimed to distinguish between benign and malignant breast nodules using an ultrasound-based intra- and peritumoral radiomics model.

Methods: This study retrospectively collected the information from 379 patients with Breast Imaging Reporting and Data System (BI-RADS) category 3–5 nodules and clear pathological diagnosis of breast nodules screened by routine ultrasound examination in the Sixth People's Hospital Affiliated to Medical College of Shanghai Jiao Tong University from January 2017 to December 2022. The largest dimension of the lesion on the 2D ultrasound image was selected to outline the area of interest which was conformally and outwardly expanded automatically by 5 mm to extract intra- and peritumor radiomics features. The included cases were randomly divided into training sets and test sets in a ratio of 7:3. The optimal features of the included models were retained by statistical and machine learning methods of dimensionality reduction, and logistic regression was used as the classifier to build an intratumoral model and a combined intratumoral-peritumoral radiomics model, respectively; through single-factor and multifactor logistic regression, the optimal features that could predict benign and malignant breast tumors were screened. The clinical and imaging models were established by selecting independent risk factors as clinical and imaging features through univariate and multifactorial logistic regression.

Results: Among 379 BI-RADS category 3–5 breast nodules, there were 124 malignant nodules and 255 benign nodules; patients were aged 14 to 88 (46.22±15.51) years, and the age differences, radiomics score, and mass diameter between the training and test sets were not statistically significant ($P>0.05$). The intra- and peritumor radiomics model had an area under the curve (AUC) of 0.840 [95% confidence interval (CI): 0.766–0.914] in the test set. The model with intra- and peritumoral ultrasound radiomics features combined with clinical features had an AUC value of 0.960 (95% CI: 0.920–0.999).

Conclusions: The nomogram, developed using intratumoral and peritumoral radiomics features combined with clinical risk features, demonstrated superior performance in distinguishing between benign and malignant BI-RADS 3–5 lesions.

Keywords: Ultrasonography; radiomics; intratumor; peritumor; breast tumors; prediction; biopsy

Submitted Mar 07, 2023. Accepted for publication Aug 28, 2023. Published online Sep 22, 2023.

doi: 10.21037/qims-23-283

View this article at: <https://dx.doi.org/10.21037/qims-23-283>

Introduction

Breast cancer, one of the major malignancies in women, is becoming more prevalent at younger ages (1) and is now the leading cause of cancer-related death in women aged 20–59 years (2). Early diagnosis and treatment are essential to improving the survival rate and quality of life of patients with breast cancer (3). Breast imaging examinations mainly include ultrasound, mammography, and magnetic resonance imaging (MRI). Mammography has advantages such as simplicity of operation, low cost, high resolution, good reproducibility, and sensitivity to calcifications (4). However, it has poor imaging efficacy for dense breasts, which are prevalent among Asian women and thus has low sensitivity and accuracy in detecting breast cancer in these women (5,6). Compared to mammography, breast ultrasound has several advantages: it is noninvasive, cost-effective, and easy to operate, while not involving radiation; however, it lacks sensitivity to detect calcification. Recent research indicates that breast ultrasound can, to some extent, compensate for the drawbacks of mammography. As a supplementary method to mammography examination, breast ultrasound can effectively detect breast cancer at an earlier stage and enhance the sensitivity and detection rate of cancer screening (7–9). Ultrasound has become one of the primary methods for the early screening of breast nodules (10). Although breast MRI yields excellent imaging effects and high sensitivity, it also has a substantially high false-positive rate (11), and limitations in usage, including patients with internal fixation surgeries or conditions such as pacemaker implantation, claustrophobia, and contrast agent allergy, restrict its implementation. As a result, MRI is primarily used for the supplemental examination of breast ultrasound and mammography, as well as screening high-risk populations, such as carriers of gene mutations (12). The American College of Radiology (ACR) Ultrasound Breast Imaging Reporting and Data System (BI-RADS) is a useful tool that can assist clinicians in correctly managing

breast nodules. However, for breast nodules categorized as BI-RADS 3 to 5, there is a broad range of possible malignancies, making it challenging to reach a differential diagnosis and to avoid unnecessary biopsy procedures (13). Radiomics can extract features from medical images that are not visually identifiable by the human eye in a high-throughput manner, and the extracted features may be associated with the tumor's heterogeneity and the biological activity of its cells (14). Reports indicate that radiomics may improve the diagnosis, prognosis, and treatment prediction of breast cancer (15). The area under the curve (AUC) values of ultrasound-based radiomic models for diagnosing breast cancer range from 0.817 to 0.943 (16–19). Other studies have shown that ultrasound-based radiomics and deep learning radiomics based on multimodal ultrasound (color Doppler or elastography) can improve the recognition ability of radiomics (20–22). Prior breast radiomics research has mainly focused on intratumoral regions. Breast cancer consists of tumor cells and stromal cells, which can cause significant changes in the peritumoral stroma. Multiregional radiomics information may be a potential means to improving the diagnosis of breast cancer (23).

This study aimed to develop a model that effectively enhances the efficacy of the preoperative diagnosis of breast cancer by extracting radiomics features from gray-scale ultrasound images of intratumoral and peritumoral regions of breast nodules and by integrating these features with clinical and ultrasound data. We present this article in accordance with the TRIPOD reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-23-283/rc>).

Methods

Statement of ethics

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was

approved by the local medical ethics committee at Sixth People's Hospital Affiliated to Medical College of Shanghai Jiao Tong University (No. 2019-027). All data used in this study underwent anonymization and did not compromise personal privacy or commercial interests. Therefore, the requirement for informed consent from all patients was waived.

Study population

Our study continuously included 379 patients who underwent routine examinations and breast lesion surgeries at our hospital between January 2017 and December 2022. Moreover, patients were required to have undergone routine ultrasound scans suggestive of BI-RADS category 3–5 breast nodules as supported by complete examination data and confirmed by surgical or puncture biopsy pathology. Patients were not enrolled if they had incomplete pathological findings, a history of radiotherapy or chemotherapy, or poor-quality ultrasound images. Enrollment additionally required all patients to have complete clinicopathological information and descriptive reports. *Figure 1A* shows the flowchart depicting the patient inclusion process.

Ultrasonography and region of interest (ROI) outlining

For ultrasonography, the patient was placed in the supine position, a multisectional radiographic scan was performed with the nipple as the center of the scan, and the lesion was observed from multiple views. The image of the largest cross-section of the node was then stored in Digital Imaging and Communications in Medicine (DICOM) format on a hard disk. Two expert diagnostic physicians with over 10 years of experience in breast ultrasound diagnosis, who were unaware of the pathological results, classified the breast nodules according to the BI-RADS. In case of controversial nodules, a consensus was reached through discussions between the 2 physicians. We collected data on age, size, color Doppler flow imaging (CDFI), and BI-RADS classification to use as variables in the clinical feature model that followed.

Images that met the inclusion criteria for an ultrasound diagnosis of a clear BI-RADS class 3 to 5 breast nodule were selected using randomization to select 70% of patients ($n=265$) to form a training set and the remaining 30% ($n=114$) to form a test set. Since the voxels were already isotropic in plane, there was no need for resampling

before extracting features. However, in order to ensure the comparability of images acquired from different scanners and with varying settings, gray-level whole-image normalization was conducted. As a result, the range was standardized to 0–600. The ROI was manually outlined along the contour of the breast mass on the ultrasound image using ITK-SNAP 3.8 software (*Figure 1B*) by a physician with more than 10 years of experience in breast ultrasound diagnosis with unknown pathology; the peritumoral ROI was a ring-like region obtained by automatically conformally expanding the boundaries of the ROI by 5 mm according to the intratumoral outline (*Figure 1B*).

Feature extraction and selection

The Pyradiomics package version 3.0.1 in Python (The Python Software Foundation) was used to calculate radiomics features. These features included first-order, shape-based (2D), and textural features and were calculated from the original ultrasound images. In addition, filtered features were calculated by filtering the original images with various types of filters, as described in the pyradiomics documentation (<https://pyradiomics.readthedocs.io/>). First-order and textural features were calculated based on the filtered ultrasound images. A total of 3,123 features were extracted from the intra- and peritumoral areas in each patient.

All extracted features were analyzed as follows. First, the intra- and interclass correlation coefficient (ICC) analyses were performed on the extracted features. Features with ICC values lower than 0.9 were excluded. Following this, the Mann-Whitney test was performed on the features, and any feature with a P value greater than 0.05 was eliminated. Features exhibiting a strong pairwise Spearman correlation ($r>0.90$) and the highest mean correlation with all other features were excluded. Next, we employed the least absolute shrinkage and selection operator (LASSO) regression model to construct the radiomics signature using the discovery dataset. LASSO shrinks all regression coefficients toward zero and sets the coefficients of many irrelevant features to zero based on the weight λ . To find the optimal λ , we used 10-fold cross-test with minimum criteria, where the final value of λ yielded the minimum cross-test error. The parameters of the retained features with nonzero coefficients were used for the regression model fitting and were combined into a radiomics signature. This was followed by obtaining a radiomics score for each patient through the linear combination of the retained

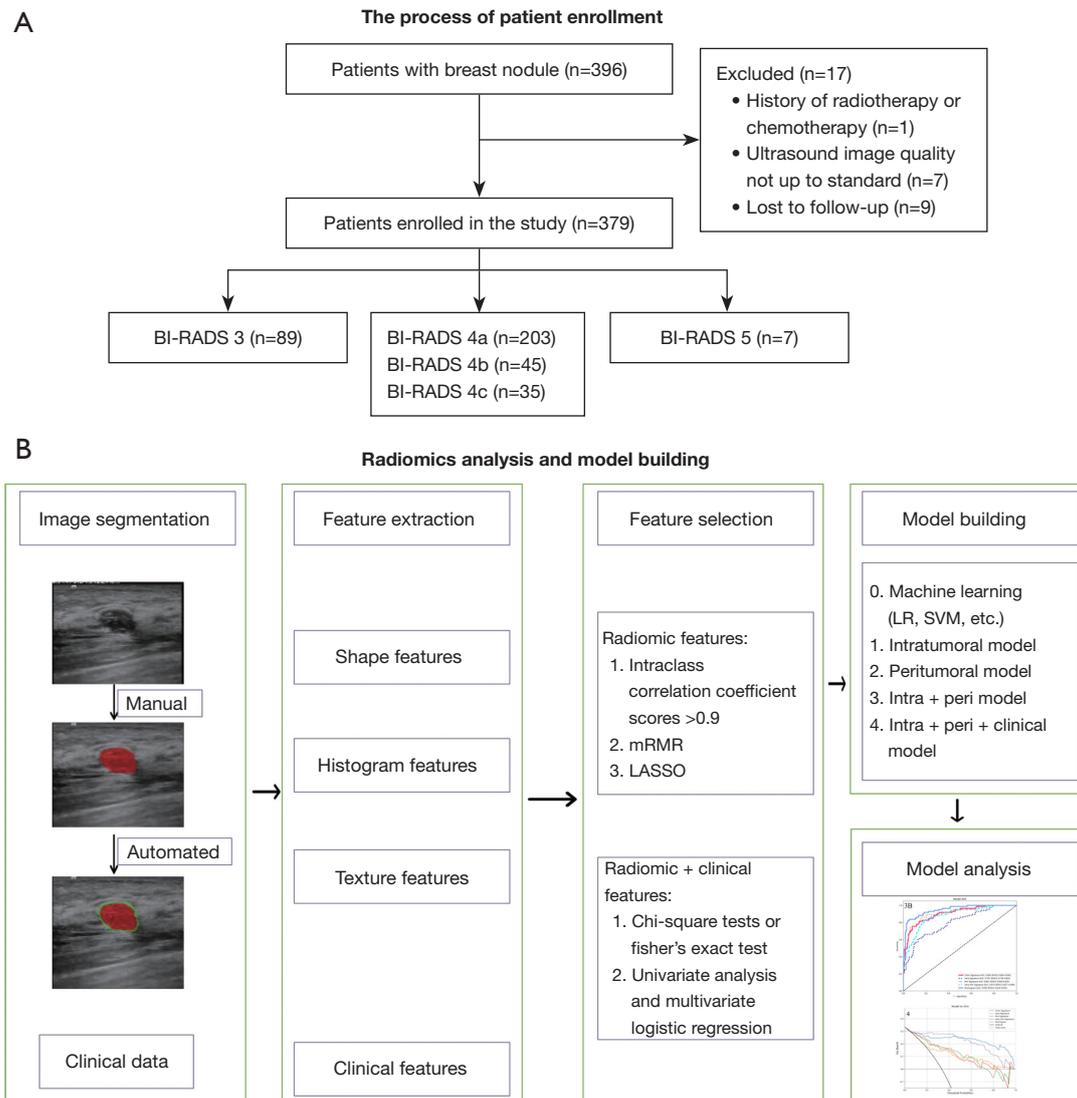


Figure 1 Study flowchart. BI-RADS, Breast Imaging Reporting and Data System; mRMR, maximum relevance-minimum redundancy; LASSO, least absolute shrinkage and selection operator; LR, logistic regression; SVM, support vector machine; intra, intratumoral features; peri, peritumoral features.

features weighed by their respective model coefficients. The Python “scikit-learn” package was used for LASSO regression modeling.

Model construction and test

We used the logistic regression machine learning model to construct the risk model using the final features. Logistic regression can be described as follows: when confronted with a regression or classification problem, a

cost function is established, and optimal model parameters are iteratively determined through optimization methods; the obtained model is then tested and validated to evaluate its performance. We adopted 5-fold cross-test to obtain the final radiomics signature. Additionally, to assess the incremental prognostic value of the radiomics signature to the clinical risk factors, we developed a radiomics nomogram based on logistic regression analysis using the test dataset. The nomogram combines the radiomics signature and clinical risk factors to calculate the breast

cancer prediction. We used the calibration curve to compare the agreement between the breast cancer prediction of the nomogram and the actual observation.

The building process for the clinical signature is similar to that of the radiomics signature. We selected the features for building the clinical signature based on baseline statistics with a P value <0.05 . We also used the same machine learning model and set the 5-fold cross-test and test cohort to be fixed for a fair comparison.

We established a radiomics nomogram by combining the radiomics signature and clinical signature and tested the diagnostic efficacy of the radiomic nomogram in the test cohort by drawing receiver operating characteristic (ROC) curves. The calibration efficiency of the nomogram was evaluated using calibration curves, and the calibration ability was assessed using the Hosmer-Lemeshow test. Additionally, we used decision curve analysis (DCA) to evaluate the clinical utility of the predictive models.

Statistical analysis

The data were statistically analyzed using R software (The R Foundation for Statistical Computing) and SPSS 26.0 (IBM Corp., Armonk, NY, USA), and the pathology results were used as the gold standard. The measures were tested for normal distribution, with the χ^2 test and 2-sample independent t -test being used to match the normal distribution. The Mann-Whitney test was used to compare nonnormally distributed continuous variables, and the χ^2 test was used to compare categorical variables. Predictive models were developed, and ROC curves were plotted to determine the accuracy, sensitivity, specificity, precision, F1 values, and AUC values of the models. Calibration curves were also plotted. Differences were considered statistically significant at $P<0.05$.

Results

Clinical characteristics

Among the 379 cases of BI-RADS category 3–5 breast nodules, postoperative pathological results revealed 255 cases of benign lesions (including 86 cases of breast adenosis, 146 cases of fibroadenoma, 4 cases of benign papillary tumors, 14 cases of intraductal papilloma, and 5 cases of granulomatous lobular mastitis) and 124 malignant lesions (including 63 cases of ductal carcinoma *in situ*, 26 cases of lobular carcinoma *in situ*, 28 cases

of invasive lobular carcinoma, and 7 cases of papillary carcinoma). The age of the patients ranged from 14 to 88 years (mean age 46.22 ± 15.51 years). Between the 2 groups, there were no statistically significant differences in age ($P=0.605$), radiomics score values ($P=0.892$), or mass diameter ($P=0.192$). Univariate logistic analysis revealed statistically significant correlations in the training set between clinical characteristics of age, mass diameter, BI-RADS classification, calcification status, and CDFI in patients with breast cancer. According to the multivariate logistic analysis, the clinical data of age, mass diameter, BI-RADS classification, calcification, CDFI, and benign and malignant breast nodules showed a significant correlation (all P values <0.001) (Table 1), which was considered as a predictive model related to BI-RADS category 3 to 5. The AUC value of the test set of the clinical model was 0.937 (95% CI: 0.885–0.989).

Feature selection and development of the model

Within the 5-mm ROI of the intratumor and peritumor, 214 radiomics features were extracted, including 36 first-order features, 28 morphological features, 150 second-order and higher-order texture features.

Of the 27 intratumoral and 37 peritumoral features screened, 15 optimal intratumoral features were selected by LASSO regression for inclusion in the intratumoral ultrasound radiomics model (Figure 2A,2B). The intratumoral ultrasound radiomics model (see Appendix 1) had an AUC value of 0.797 (95% CI: 0.739–0.855) in the training set and an AUC value of 0.780 (95% CI: 0.685–0.875) in the test set. The performance of radiomics model is reported in Table 2.

Of the 27 intratumoral and 37 peritumoral features screened, 18 optimal intratumoral features and 21 optimal peritumoral features were selected by LASSO regression for inclusion in the intratumoral and peritumoral ultrasound radiomics model (Figure 2C,2D). The AUC value of the training set of the intratumoral and peritumoral ultrasound radiomics model was 0.910 (95% CI: 0.871–0.949), and the AUC value of the test set of the intratumoral and peritumoral ultrasound radiomics model was 0.840 (95% CI: 0.766–0.915).

Combined model construction and test

We constructed a comprehensive model incorporating intratumoral, peritumoral, and clinical risk factors to

Table 1 Clinical information for the training and test sets

Clinical features	Training set				Test set			
	All	Benign	Malignant	P value	All	Benign	Malignant	P value
Maximum diameter (mm)	17.83±10.71	15.21±7.07	23.68±14.55	<0.001	17.68±7.61	15.87±7.00	20.78±7.69	<0.001
Age (years)	45.95±15.78	40.49±13.40	58.13±13.81	<0.001	46.85±14.90	41.86±14.36	55.40±11.66	<0.001
Calcification				<0.001				<0.001
None	192 (0.72)	152 (0.83)	40 (0.49)		96 (0.84)	67 (0.93)	29 (0.69)	
Yes	73 (0.28)	31 (0.17)	42 (0.51)		18 (0.16)	5 (0.07)	13 (0.31)	
CDFI				<0.001				<0.001
None	162 (0.61)	136 (0.74)	26 (0.32)		83 (0.73)	62 (0.86)	21 (0.50)	
Yes	103 (0.39)	47 (0.26)	56 (0.68)		31 (0.27)	10 (0.14)	21 (0.50)	
BI-RADS				<0.001				<0.001
3–4a	220 (0.83)	177 (0.97)	43 (0.52)		72 (0.63)	64 (0.89)	8 (0.19)	
4b–5	45 (0.17)	6 (0.03)	39 (0.48)		42 (0.37)	8 (0.11)	34 (0.81)	

Continuous data are presented as the mean ± standard deviation; categorical data are presented as number (ratio). BI-RADS category was determined by an ultrasound diagnostic physician. CDFI, color Doppler flow imaging; BI-RADS, Breast Imaging Reporting and Data System.

develop a nomogram (shown in [Appendix 1](#)) for predicting the benign or malignant nature of BI-RADS category 3–5 breast nodules ([Figure 3A](#)). The intratumoral and peritumoral ultrasound radiomics features combined with clinical features had an AUC value of 0.950 (95% CI: 0.924–0.976) in the training set column line graph model and an AUC value of 0.960 (95% CI: 0.920–0.999) in the test set column line graph model ([Figure 3B,3C](#)). The calibration curve of intratumoral and peritumoral ultrasound radiomics features combined with clinical features shows that it had good calibration in both the training and test sets ([Figure 3D,3E](#)). The DCA indicated that the nomogram-based model incorporating intra- and peritumoral ultrasound radiomics and clinical features had a higher overall net benefit than did the other 2 models across the majority of the range of reasonable threshold probabilities, which means that the nomogram was useful for discriminating between benign and malignant lesion in BI-RADS category 3–5 nodules ([Figure 4](#)). The model of intratumoral and peritumoral ultrasound radiomics combined with clinical features was effective in reducing the number of unnecessary puncture biopsies in breast nodules, especially in BI-RADS category 4. When the nomogram score was greater than 0.5, the nodule had a high degree of malignancy and required be biopsy; when the nomogram score was less than 0.5, the nodule had a low degree of

malignancy and could be followed up only.

Discussion

Although the use of BI-RADS criteria has improved the accuracy of the results, it is highly subjective and dependent on the diagnostic ability of the examining physician. There is often some overlap between the ultrasound presentation of benign and malignant breast tumors (24,25). Moreover, medical images often contain a wealth of information that reflects the biological behavior of the tumor—information that is difficult for the human eye to recognize. With radiomics, more objective and precise quantitative features can be extracted from medical images through high-throughput calculations, and objective and quantitative data, such as tumor heterogeneity, can be used to describe tumor phenotypes and provide valuable information for tumor diagnosis and treatment (26). In the study by Hong *et al.* (27), the radiomics feature model they developed for diagnosing BI-RADS category 4–5 breast nodules achieved an AUC value of 0.886 on the test set. Similarly, in the study by Romeo *et al.* (19), the combination of radiomics and machine learning yielded an AUC value of 0.82. In our study, we used LASSO regression analysis to select 2 morphological features, 2 first-order histogram features, and 11 texture features from the intratumoral radiomics

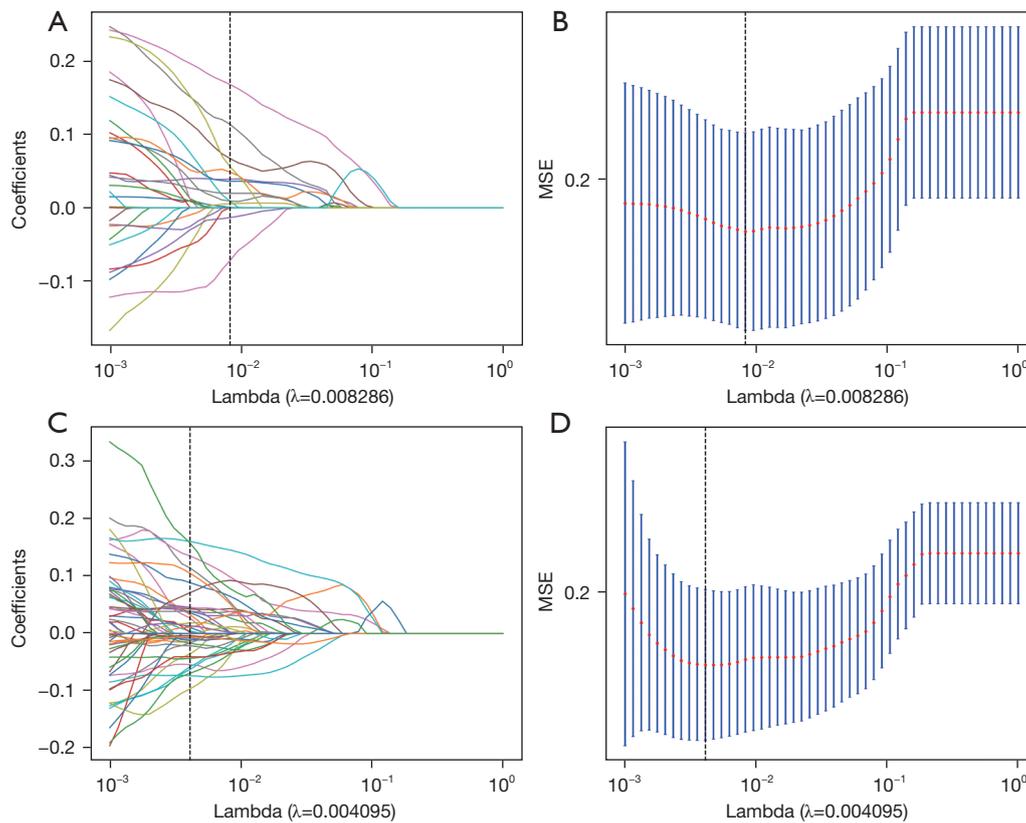


Figure 2 The feature selection process. The parameters are screened with λ being used to minimize the mean error. The LASSO logistic regression model was used to select radiomics features. (A,B) The feature selection process of intratumoral features. (C,D) The feature selection process of intratumoral and peritumoral features combined. MSE, meansquared error; LASSO, least absolute shrinkage and selection operator.

Table 2 Diagnostic performance of model

Model	Set	AUC	Accuracy	Sensitivity	Specificity	F1
Intra-radiomics	Training	0.797	0.785	0.534	0.905	0.73
	Test	0.780	0.754	0.526	0.868	0.667
Intra + peri-radiomics	Training	0.910	0.868	0.721	0.938	0.779
	Test	0.840	0.745	0.500	0.868	0.567
Intra + peri + clinical	Training	0.950	0.913	0.791	0.972	0.855
	Test	0.960	0.921	0.816	0.974	0.873

AUC, area under the curve; intra, intratumoral features; peri, peritumoral features.

features and incorporated them into the intratumoral ultrasound radiomics model. The AUC value in the training set and test set was 0.800 (95% CI: 0.739–0.855) and 0.780 (95% CI: 0.6854–0.875), respectively, slightly lower than values in the 2 studies mentioned above (19,27).

Previous ultrasound radiomics studies of breast tumors have focused mostly on intratumoral and not peritumoral radiomics features; however, there may be some important biological information in the peritumoral region of breast cancer, such as interstitial reaction, angiogenesis, and

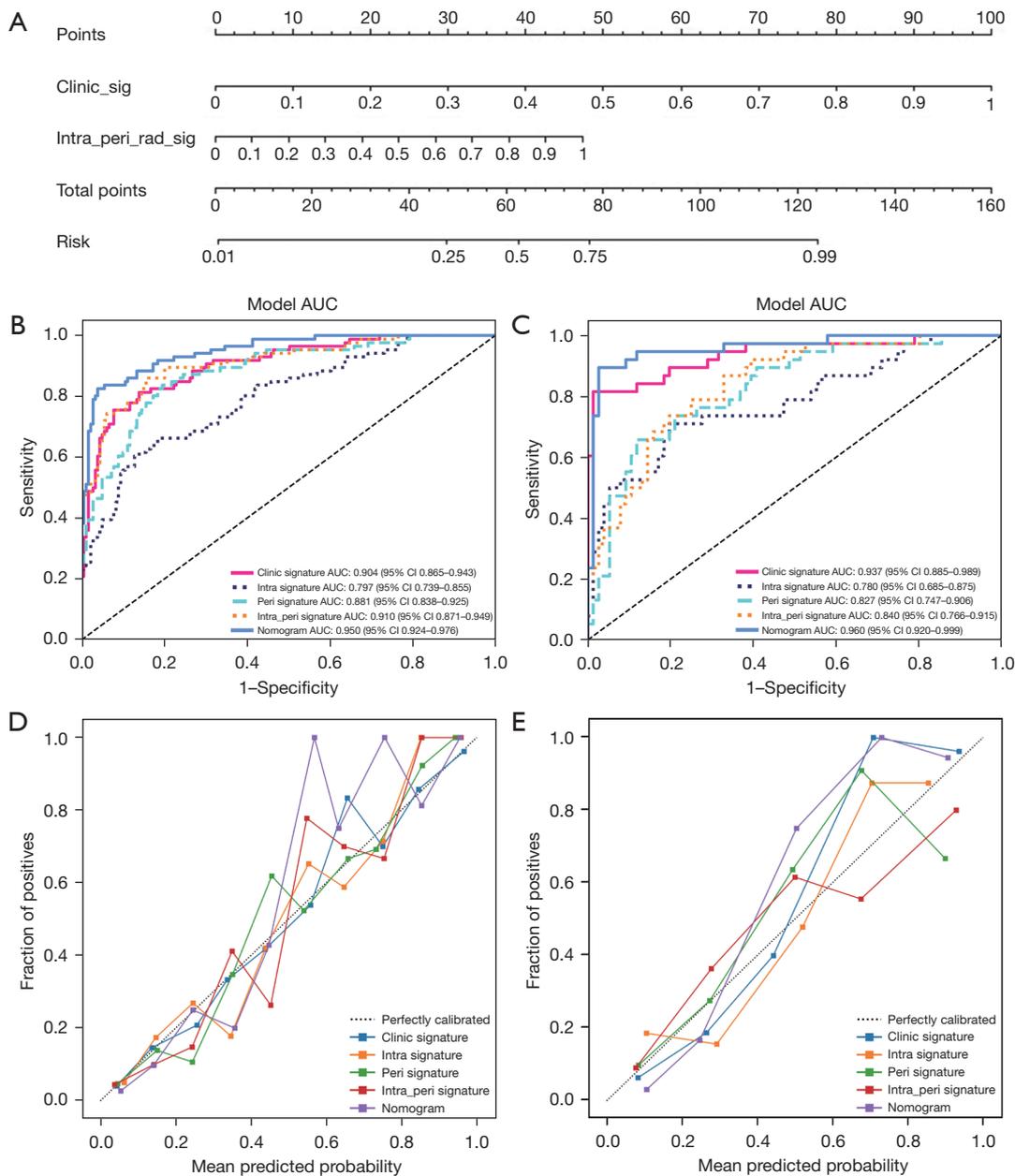


Figure 3 A radiomics nomogram (A) was developed in which the intratumoral peritumoral ultrasound radiomics were combined with clinical and ultrasound features. The receiver operating characteristics of the radiomics nomogram, intratumoral model (intra signature), and intratumoral combined peritumoral model (intra_peri signature) in the differential diagnosis of breast nodules were compared between the training set (B) and the test set (C). The radiomics nomogram calibration curves for the training set (D) and test set (E). intra, intratumoral features; peri, peritumoral features; AUC, area under the curve; CI, confidence interval.

peritumoral infiltration of lymphatic vessels and blood vessels (28,29). We attempted to combine intratumoral and peritumoral radiomics features and incorporated 18 intratumoral radiomics features and 21 peritumoral

radiomics features into our diagnostic model, whose AUC values improved from 0.797 to 0.904 in the training set and from 0.780 to 0.937 in the test set, with a statistically significant difference ($P=0.04$) (30). The combined

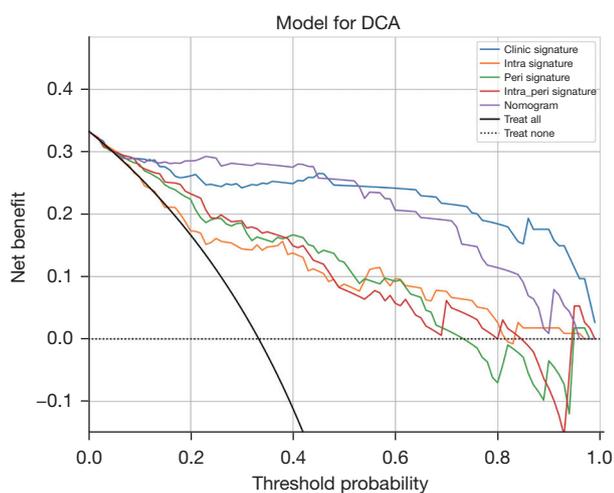


Figure 4 Analysis of decision curves for each model in the test sets. DCA, decision curve analysis; intra, intratumoral features; peri, peritumoral features.

intratumoral and peritumoral quantitative ultrasound features achieved an AUC of 0.94 in identifying benign and malignant breast lesions, compared with an AUC of 0.83 when intratumoral quantitative ultrasound features were used alone. The diagnostic model developed in this study consisting of combined intratumoral and peritumoral features was superior to the radiomics model developed with intratumoral features alone. This is possibly because some BI-RADS category 3–5 nodules are difficult to identify through intratumoral radiomics features but can be identified through their peritumoral radiomics features. Therefore, peritumoral radiomic features may be potentially helpful in situations where intratumoral radiomic features alone cannot provide accurate discrimination. Furthermore, other studies have found that radiomic features based on X-ray, contrast-enhanced X-ray, MRI, and contrast-enhanced MRI in both the intratumoral and peritumoral regions are associated with the classification of benign and malignant breast lesions (31), lymph node metastasis (32–34), estrogen receptor and progesterone receptor expression levels (35), human epidermal growth factor 2 (HER-2) and Ki-67 status (36), neoadjuvant chemotherapy effect (37), and the molecular subtype of breast cancer (38).

Our study found that many clinical factors were associated with benign and malignant breast lesions, including patient age, mass diameter, BI-RADS classification, calcification status, and CDFI. Therefore, we developed a nomogram incorporating both intratumoral

and peritumoral radiomics features and clinical features that demonstrated excellent diagnostic power for the differential diagnosis of benign and malignant BI-RADS category 3–5 breast nodules. The proposed model nomogram achieved an AUC of 0.960, which was 0.12 higher than the AUC of 0.840 achieved by the intra- and peritumoral radiomics model (0.840); meanwhile, the sensitivity was significantly increased from 0.500 to 0.816 and the specificity from 0.868 to 0.974, an improvement which may help to mitigate the incidence of false negatives and false positives (39). In our study, the high sensitivity and specificity observed might be attributed to the limited number of breast malignant tumor types included as well as the lack of racial diversity in our patient population. It has been suggested by some researchers that artificial intelligence models perform better on images that are similar to those in the training dataset, and caution must be taken when encountering potentially underrepresented subgroups in the training dataset, such as those related to racial groups and device vendors (40). In a previous study, intratumoral ultrasound radiomics were combined with BI-RADS classification to predict benign and malignant breast lesions (39). Their results showed that the model combining the radiomics score and the BI-RADS category showed better ability to discriminate between malignant and benign breast lesions (AUC: 0.928) than did the radiomics score model (AUC: 0.857) and the BI-RADS category model (AUC: 0.864). In Hong *et al.*'s study (27), a diagnostic model combining extracted intratumoral ultrasound radiomics features with patient age, BI-RADS category, and maximum lesion diameter showed a good ability to discriminate between malignant and benign breast lesions (AUC: 0.937). However, the diagnostic performance of the combined models used by Luo *et al.* (39) and Hong *et al.* (27) was worse than that of our model combining intratumoral and peritumoral radiomics feature with clinical features, demonstrating the superiority of our diagnostic model. The model constructed by Hong *et al.* (27), which was based on 496 BI-RADS category 4 or 5 lesions and consists of radiomics score, age, ultrasound-detected lesion diameter, and BI-RADS category, achieved an AUC of 0.937 (95% CI: 0.893–0.965) in the test cohort. Meanwhile, Luo *et al.*'s model (27), which was based on 315 BI-RADS category 4 or 5 lesions and consists of radiomics score and BI-RADS category, yielded an AUC of 0.928 (95% CI: 0.876–0.980). Of course, all the aforementioned studies have one limitation in common: they are retrospective studies. The true diagnostic efficacy of these models in clinical practice lacks test from large-scale, multi-institutional

prospective trials conducted by independent third parties.

There are also some limitations to the present study. First, we employed a single-center study design and a small sample size, which might have resulted in reduced model stability. Additionally, the imbalance in the number of included benign and malignant cases may have some impact on the diagnostic performance of our models. Second, our study was also retrospective in nature; therefore, the possibility of selection bias cannot be excluded. Future research should consist of a prospective study to confirm our models. Third, both the intratumoral and peritumoral ROIs were outlined at the 2D level only, and the 3D features of the tumor were ignored. Fourth, 5 mm was chosen as the peritumoral ROI, and the peritumoral features beyond 5 mm from the intratumoral ROI were not extracted, which could have missed some of the radiomics feature information of the peritumor. Finally, some of the clinical features (such as BI-RADS category, ROI segmentation) in this study are semiquantitative features, and the subjectivity of the evaluators might have influenced their results.

In summary, a combined diagnostic model based on intra- and peritumoral radiomics features combined with clinical and imaging features can more accurately identify benign and malignant BI-RADS class 3–5 breasts nodules, providing guidance to clinicians in their decision-making and thereby reducing the rate of missed diagnoses and avoiding unnecessary biopsies.

Acknowledgments

Funding: The study was supported in part by Shanghai Municipal Health Commission, Shanghai, China (No. 20204Y0343).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-23-283/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-23-283/coif>). 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. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the local medical ethics committee at the Sixth People's Hospital Affiliated to Medical College of Shanghai Jiao Tong University (No. 2019-027). All data used in this study underwent anonymization and did not compromise personal privacy or commercial interests. As a result, the requirement for informed consent from all patients 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. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209-49.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin* 2021;71:7-33.
3. Zhong LC, Yang T, Gu LP, Ma F. The diagnostic performance of shear wave velocity ratio for the differential diagnosis of benign and malignant breast lesions: Compared with VTQ, and mammography. *Clin Hemorheol Microcirc* 2021;77:123-31.
4. Johnson K, Lång K, Ikeda DM, Åkesson A, Andersson I, Zackrisson S. Interval Breast Cancer Rates and Tumor Characteristics in the Prospective Population-based Malmö Breast Tomosynthesis Screening Trial. *Radiology* 2021;299:559-67.
5. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, Jong RA, Hislop G, Chiarelli A, Minkin S, Yaffe MJ. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227-36.
6. Geisel J, Raghu M, Hooley R. The Role of Ultrasound in Breast Cancer Screening: The Case for and Against Ultrasound. *Semin Ultrasound CT MR* 2018;39:25-34.
7. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, Böhm-Vélez M, Mahoney MC, Evans WP

- 3rd, Larsen LH, Morton MJ, Mendelson EB, Farria DM, Cormack JB, Marques HS, Adams A, Yeh NM, Gabrielli G; . Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 2012;307:1394-404.
8. Ohuchi N, Suzuki A, Sobue T, Kawai M, Yamamoto S, Zheng YF, Shiono YN, Saito H, Kuriyama S, Tohno E, Endo T, Fukao A, Tsuji I, Yamaguchi T, Ohashi Y, Fukuda M, Ishida T; J-START investigator groups. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. *Lancet* 2016;387:341-8.
 9. Tagliafico AS, Calabrese M, Mariscotti G, Durando M, Tosto S, Monetti F, Airaldi S, Bignotti B, Nori J, Bagni A, Signori A, Sormani MP, Houssami N. Adjunct Screening With Tomosynthesis or Ultrasound in Women With Mammography-Negative Dense Breasts: Interim Report of a Prospective Comparative Trial. *J Clin Oncol* 2016;34:1882-8.
 10. Drukker K, Giger ML, Horsch K, Kupinski MA, Vyborny CJ, Mendelson EB. Computerized lesion detection on breast ultrasound. *Med Phys* 2002;29:1438-46.
 11. Kuhl CK, Keulers A, Strobel K, Schneider H, Gaisa N, Schrading S. Not all false positive diagnoses are equal: On the prognostic implications of false-positive diagnoses made in breast MRI versus in mammography / digital tomosynthesis screening. *Breast Cancer Res* 2018;20:13.
 12. Saslow D, Boetes C, Burke W, Harms S, Leach MO, Lehman CD, Morris E, Pisano E, Schnall M, Sener S, Smith RA, Warner E, Yaffe M, Andrews KS, Russell CA; American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin* 2007;57:75-89.
 13. D'Orsi CJ, Sickles E, Mendelson E, Morris EA. ACR BI-RADS® Atlas: Breast Imaging Reporting and Data System. Reston, VA, USA: American College of Radiology; 2013. Available online: <https://www.scienceopen.com/book?vid=ddb508dc-e150-4b1c-aad8-1be1c031edd8>
 14. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics. *Phys Med* 2017;38:122-39.
 15. Conti A, Duggento A, Indovina I, Guerrisi M, Toschi N. Radiomics in breast cancer classification and prediction. *Semin Cancer Biol* 2021;72:238-50.
 16. Fleury E, Marcomini K. Performance of machine learning software to classify breast lesions using BI-RADS radiomic features on ultrasound images. *Eur Radiol Exp* 2019;3:34.
 17. Fleury EFC, Marcomini K. Impact of radiomics on the breast ultrasound radiologist's clinical practice: From lumpologist to data wrangler. *Eur J Radiol* 2020;131:109197.
 18. Li J, Bu Y, Lu S, Pang H, Luo C, Liu Y, Qian L. Development of a Deep Learning-Based Model for Diagnosing Breast Nodules With Ultrasound. *J Ultrasound Med* 2021;40:513-20.
 19. Romeo V, Cuocolo R, Apolito R, Stanzione A, Ventimiglia A, Vitale A, Verde F, Accurso A, Amitrano M, Insabato L, Gencarelli A, Buonocore R, Argenzio MR, Cascone AM, Imbriaco M, Maurea S, Brunetti A. Clinical value of radiomics and machine learning in breast ultrasound: a multicenter study for differential diagnosis of benign and malignant lesions. *Eur Radiol* 2021;31:9511-9.
 20. Qian X, Pei J, Zheng H, Xie X, Yan L, Zhang H, Han C, Gao X, Zhang H, Zheng W, Sun Q, Lu L, Shung KK. Prospective assessment of breast cancer risk from multimodal multiview ultrasound images via clinically applicable deep learning. *Nat Biomed Eng* 2021;5:522-32.
 21. Shen Y, Shamout FE, Oliver JR, Witowski J, Kannan K, Park J, et al. Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams. *Nat Commun* 2021;12:5645.
 22. Zhang Q, Song S, Xiao Y, Chen S, Shi J, Zheng H. Dual-mode artificially-intelligent diagnosis of breast tumours in shear-wave elastography and B-mode ultrasound using deep polynomial networks. *Med Eng Phys* 2019;64:1-6.
 23. Mao Y, Keller ET, Garfield DH, Shen K, Wang J. Stromal cells in tumor microenvironment and breast cancer. *Cancer Metastasis Rev* 2013;32:303-15.
 24. Evans A, Whelehan P, Thomson K, Brauer K, Jordan L, Purdie C, McLean D, Baker L, Vinnicombe S, Thompson A. Differentiating benign from malignant solid breast masses: value of shear wave elastography according to lesion stiffness combined with greyscale ultrasound according to BI-RADS classification. *Br J Cancer* 2012;107:224-9.
 25. Youk JH, Kwak JY, Lee E, Son EJ, Kim JA. Grayscale Ultrasound Radiomic Features and Shear-Wave Elastography Radiomic Features in Benign and Malignant Breast Masses. *Ultraschall Med* 2020;41:390-6.
 26. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 2016;278:563-77.
 27. Hong ZL, Chen S, Peng XR, Li JW, Yang JC, Wu SS. Nomograms for prediction of breast cancer in breast

- imaging reporting and data system (BI-RADS) ultrasound category 4 or 5 lesions: A single-center retrospective study based on radiomics features. *Front Oncol* 2022;12:894476.
28. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, Plecha D, Madabhushi A. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res* 2017;19:57.
 29. Yu H, Meng X, Chen H, Han X, Fan J, Gao W, Du L, Chen Y, Wang Y, Liu X, Zhang L, Ma G, Yang J. Correlation Between Mammographic Radiomics Features and the Level of Tumor-Infiltrating Lymphocytes in Patients With Triple-Negative Breast Cancer. *Front Oncol* 2020;10:412.
 30. Klimonda Z, Karwat P, Dobruch-Sobczak K, Piotrkowska-Wróblewska H, Litniewski J. Breast-lesions characterization using Quantitative Ultrasound features of peritumoral tissue. *Sci Rep* 2019;9:7963.
 31. Wang S, Sun Y, Mao N, Duan S, Li Q, Li R, Jiang T, Wang Z, Xie H, Gu Y. Incorporating the clinical and radiomics features of contrast-enhanced mammography to classify breast lesions: a retrospective study. *Quant Imaging Med Surg* 2021;11:4418-30.
 32. Cheng Y, Xu S, Wang H, Wang X, Niu S, Luo Y, Zhao N. Intra- and peri-tumoral radiomics for predicting the sentinel lymph node metastasis in breast cancer based on preoperative mammography and MRI. *Front Oncol* 2022;12:1047572.
 33. Lin F, Li Q, Wang Z, Shi Y, Ma H, Zhang H, Zhang K, Yang P, Zhang R, Duan S, Gu Y, Mao N, Xie H. Intratumoral and peritumoral radiomics for preoperatively predicting the axillary non-sentinel lymph node metastasis in breast cancer on the basis of contrast-enhanced mammography: a multicenter study. *Br J Radiol* 2023. [Epub ahead of print]. doi: 10.1259/bjr.20220068.
 34. Sun Q, Lin X, Zhao Y, Li L, Yan K, Liang D, Sun D, Li ZC. Deep Learning vs. Radiomics for Predicting Axillary Lymph Node Metastasis of Breast Cancer Using Ultrasound Images: Don't Forget the Peritumoral Region. *Front Oncol* 2020;10:53.
 35. Zhong S, Wang F, Wang Z, Zhou M, Li C, Yin J. Multiregional Radiomic Signatures Based on Functional Parametric Maps from DCE-MRI for Preoperative Identification of Estrogen Receptor and Progesterone Receptor Status in Breast Cancer. *Diagnostics (Basel)* 2022;12:2558.
 36. Li C, Song L, Yin J. Intratumoral and Peritumoral Radiomics Based on Functional Parametric Maps from Breast DCE-MRI for Prediction of HER-2 and Ki-67 Status. *J Magn Reson Imaging* 2021;54:703-14.
 37. Mao N, Shi Y, Lian C, Wang Z, Zhang K, Xie H, Zhang H, Chen Q, Cheng G, Xu C, Dai Y. Intratumoral and peritumoral radiomics for preoperative prediction of neoadjuvant chemotherapy effect in breast cancer based on contrast-enhanced spectral mammography. *Eur Radiol* 2022;32:3207-19.
 38. Niu S, Jiang W, Zhao N, Jiang T, Dong Y, Luo Y, Yu T, Jiang X. Intra- and peritumoral radiomics on assessment of breast cancer molecular subtypes based on mammography and MRI. *J Cancer Res Clin Oncol* 2022;148:97-106.
 39. Luo WQ, Huang QX, Huang XW, Hu HT, Zeng FQ, Wang W. Predicting Breast Cancer in Breast Imaging Reporting and Data System (BI-RADS) Ultrasound Category 4 or 5 Lesions: A Nomogram Combining Radiomics and BI-RADS. *Sci Rep* 2019;9:11921.
 40. Bhowmik A, Eskreis-Winkler S. Deep learning in breast imaging. *BJR Open* 2022;4:20210060.

Cite this article as: Zhong L, Shi L, Zhou L, Liu X, Gu L, Bai W. Development of a nomogram-based model combining intra- and peritumoral ultrasound radiomics with clinical features for differentiating benign from malignant in Breast Imaging Reporting and Data System category 3–5 nodules. *Quant Imaging Med Surg* 2023;13(10):6899–6910. doi: 10.21037/qims-23-283

Appendix 1

1. Rad-score (intra-tumoral ultrasound Radiomics model) = $0.3245283018867949 + 0.000366 * \text{intra_original_firstorder_Minimum} - 0.000609 * \text{intra_original_firstorder_TotalEnergy} + 0.038934 * \text{intra_original_gldm_Idmn} + 0.009054 * \text{intra_original_gldm_MCC} + 0.008807 * \text{intra_original_glrlm_LongRunHighGrayLevelEmphasis} + 0.035892 * \text{intra_original_glrlm_LowGrayLevelRunEmphasis} + 0.048566 * \text{intra_original_glrlm_RunEntropy} + 0.067070 * \text{intra_original_glszm_GrayLevelNonUniformity} - 0.072016 * \text{intra_original_glszm_LargeAreaHighGrayLevelEmphasis} + 0.019576 * \text{intra_original_glszm_SizeZoneNonUniformity} + 0.005535 * \text{intra_original_glszm_SmallAreaLowGrayLevelEmphasis} - 0.013511 * \text{intra_original_ngtdm_Contrast} + 0.168121 * \text{intra_original_shape_Elongation} + 0.113672 * \text{intra_original_shape_Maximum3DDiameter} + 0.055753 * \text{intra_original_shape_SurfaceVolumeRatio}$.

2. Rad-score (intra-tumoral combined peri-tumoral ultrasound Radiomics model) = $0.3245283018867948 + 0.010431 * \text{intra_original_firstorder_Minimum} - 0.012548 * \text{intra_original_firstorder_TotalEnergy} + 0.044222 * \text{intra_original_gldm_Contrast} + 0.038485 * \text{intra_original_gldm_Idmn} + 0.026819 * \text{intra_original_gldm_Imc1} - 0.035756 * \text{intra_original_gldm_GrayLevelNonUniformity} + 0.004343 * \text{intra_original_glrlm_LongRunHighGrayLevelEmphasis} + 0.036033 * \text{intra_original_glrlm_LowGrayLevelRunEmphasis} + 0.032856 * \text{intra_original_glrlm_RunEntropy} + 0.001761 * \text{intra_original_glrlm_RunPercentage} + 0.045726 * \text{intra_original_glrlm_ShortRunEmphasis} - 0.054379 * \text{intra_original_glszm_LargeAreaHighGrayLevelEmphasis} + 0.038748 * \text{intra_original_glszm_SizeZoneNonUniformity} - 0.057946 * \text{intra_original_glszm_ZoneEntropy} - 0.043282 * \text{intra_original_ngtdm_Busyness} - 0.007698 * \text{intra_original_ngtdm_Contrast} + 0.042842 * \text{intra_original_ngtdm_Strength} + 0.045331 * \text{intra_original_shape_SurfaceVolumeRatio} + 0.114245 * \text{peri_original_firstorder_10Percentile} - 0.096648 * \text{peri_original_firstorder_Maximum} - 0.062812 * \text{peri_original_firstorder_Minimum} - 0.011152 * \text{peri_original_firstorder_RobustMeanAbsoluteDeviation} - 0.017004 * \text{peri_original_gldm_ClusterProminence} + 0.135194 * \text{peri_original_gldm_Idn} - 0.021447 * \text{peri_original_gldm_Imc1} + 0.018190 * \text{peri_original_gldm_MaximumProbability} - 0.073709 * \text{peri_original_gldm_DependenceVariance} + 0.010953 * \text{peri_original_glrlm_GrayLevelNonUniformity} - 0.069274 * \text{peri_original_glrlm_LongRunLowGrayLevelEmphasis} + 0.071068 * \text{peri_original_glszm_GrayLevelNonUniformity} - 0.016455 * \text{peri_original_glszm_GrayLevelVariance} + 0.088708 * \text{peri_original_glszm_LowGrayLevelZoneEmphasis} + 0.105085 * \text{peri_original_glszm_ZoneEntropy} - 0.040601 * \text{peri_original_glszm_ZoneVariance} + 0.013239 * \text{peri_original_ngtdm_Busyness} - 0.003644 * \text{peri_original_ngtdm_Coarseness} + 0.160328 * \text{peri_original_shape_Elongation} - 0.006252 * \text{peri_original_shape_SurfaceVolumeRatio} + 0.157836 * \text{peri_original_shape_VoxelVolume}$