scientific reports

OPEN

The value of intratumoral and peritumoral ultrasound radiomics model constructed using multiple machine learning algorithms for non-mass breast cancer

Jinhui Liu[1,7], Junjun Chen[1,7], Lin Qiu[3], Ruoxin Li[1,6], Yuning Li[1,4], Ting Li[5,6] & Xiaoling Leng[1,2✉]

To investigate the diagnostic capability of multiple machine learning algorithms combined with intratumoral and peritumoral ultrasound radiomics models for non-massive breast cancer in dense breast backgrounds. Manual segmentation of ultrasound images was performed to define the intratumoral region of interest (ROI), and five peritumoral ROIs were generated by extending the contours by 1 to 5 mm. A total of 851 radiomics features were extracted from these regions and filtered using statistical methods. Thirteen machine learning algorithms were employed to create radiomics models for the intratumoral and peritumoral areas. The best model was combined with clinical ultrasound predictive factors to form a joint model, which was evaluated using ROC curves, calibration curves, and decision curve analysis (DCA).Based on this model, a nomogram was developed, demonstrating high predictive performance, with C-index values of 0.982 and 0.978.The model incorporating the intratumoral and peritumoral 2 mm regions outperformed other models, indicating its effectiveness in distinguishing between benign and malignant breast lesions. This study concludes that ultrasound imaging, particularly in the intratumoral and peritumoral 2 mm regions, has significant potential for diagnosing non-massive breast cancer, and the nomogram can assist clinical decision-making.

Keywords Ultrasound, Non-mass breast lesions, Non-massive breast cancer, Radiomics, Machine learning

Non-mass breast lesions (NML) account for 9.21% of breast abnormalities[1]. In conventional ultrasound examinations, these lesions lack clear boundaries and do not show significant space-occupying effects in two or more scanning planes. They typically appear as localized hypoechoic areas on ultrasound images, without distinct shape or border[2]. In the Breast Imaging Reporting and Data System (BI-RADS), diagnostic criteria for breast NML are lacking, which increases the risk of missed or incorrect diagnoses[3]. Based on their conventional ultrasound features, Wang et al.[4] classified breast NML into four types: hypoechoic areas within the breast, hypoechoic areas with microcalcifications, architectural distortion, and solid echoes within the ducts. However, dense breast tissue is more common among Asian women. In BI-RADS, breast tissue is categorized into[5]: fatty, scattered fibroglandular, heterogeneously dense (where masses may appear indistinct), and extremely dense (where the high density of breast tissue significantly reduces the sensitivity of mammography). The latter two types are considered dense breasts. According to the ultrasound classification proposed by Wu Lin et al.[6], dense breasts are characterized by a duct diameter sum that is less than half of the breast's thickness, with breast parenchyma predominantly composed of fibrous or fatty tissue, exhibiting high and uniform echo intensity with heterogeneous internal echoes. On ultrasound images, dense breasts predominantly show as areas of high or intense echo, with the parenchyma mostly consisting of proliferating fibrous tissue. The mammary glands may

[1]Department of Ultrasonography, The Tenth Affiliated Hospital of Southern Medical University (Dongguan People's Hospital), Dongguan 523059, Guangdong, China. [2]Department of Ultrasonography, The Affiliated Tumor Hospital of Xinjiang Medical University, Urumqi 830011, Xinjiang, China. [3]Department of Ultrasonography, Chenzhou First People's Hospital, Chenzhou 423000, Hunan, China. [4]Southern Medical University, Guangzhou 510515, Guangdong, China. [5]Department of Ultrasonography, Shenzhen Longhua District People's Hospital, Shenzhen 518109, Guangdong, China. [6]Guangdong Medical University, Zhanjiang 524023, Guangdong, China. [7]Jinhui Liu and Junjun Chen contributed equally to this work. ✉email: lengxiaoling1206@163.com

undergo atrophy, and ducts may become small or disappear[7]. The dense glandular tissue may obscure lesions, making the detection and differentiation between benign and malignant breast NML in the context of dense breast tissue particularly challenging.

In recent years, with the deepening of research, radiomics has made significant progress in the diagnosis of breast tumors, lymph node metastasis assessment, and prognosis prediction[8,9]. By precisely annotating medical images, delineating lesions, and applying advanced computational methods, this technology can extract quantitative data from a large volume of medical images[10,11]. However, most published tumor radiomics studies primarily focus on extracting radiomic features from the primary tumor site, neglecting the peritumoral radiomic features. In 1989, the "seed and soil" hypothesis was proposed[12], suggesting that the tumor microenvironment, which includes tumor cells and surrounding immune cells, plays a critical role in tumor progression and metastasis through complex interactions that allow tumor cells to evade immune responses[13]. The study by Tan et al.[14] demonstrated that a model combining texture analysis with conventional Magnetic Resonance Imaging(MRI) features had superior diagnostic performance in differentiating benign and malignant breast NML compared to using texture analysis or conventional MRI features alone. This suggests that texture analysis can quantify tumor heterogeneity and provide added value in distinguishing between benign and malignant breast NML[14].

Advances in radiomics are rapidly transforming breast cancer diagnosis and treatment, particularly by enabling deeper insights into tumor microenvironment heterogeneity. Although mammography remains a primary screening tool, its dependence on morphological characteristics limits its specificity in dense breast tissue, often resulting in a high number of unnecessary biopsies[15]. Emerging evidence suggests that peritumoral texture analysis may help address this limitation: mammographic studies have shown that wavelet-based features of stromal heterogeneity can predict malignancy risk[16], while three-compartment compositional imaging (3CB) reduces false positives by 22% through the analysis of fat-to-fibroglandular texture ratios[17]. These findings are consistent with preclinical research on tumor-adipocyte interactions[17], which demonstrate that gap junction-mediated stromal remodeling generates detectable microenvironmental signatures. However, research on the application of ultrasound-based radiomics for diagnosing non-mass breast cancer remains scarce. Most non-mass lesions (NMLs) in the breast are abnormal, and conventional ultrasound often yields high sensitivity but low specificity, leading to a relatively high rate of missed or incorrect diagnoses for non-mass breast cancer[6]. Building upon these developments, we hypothesize that ultrasound radiomics—particularly peri_5mm wavelet features—can detect early stromal reprogramming in non-mass lesions, thereby addressing a crucial diagnostic gap in populations with dense breast tissue.

Methods
Pathological selection and general information
A total of 851 cases of breast NWL in a dense breast background, detected by conventional ultrasound, were included in this study between May 30, 2018, and January 20, 2024, at Dongguan People's Hospital. Detailed information on the enrollment procedure is displayed in Fig. 1. Inclusion criteria: (I) Diagnosed as breast NML by ultrasound; (II) Diagnosed with dense breast tissue by ultrasound or mammography; (III) Patients who voluntarily underwent core needle biopsy, vacuum-assisted biopsy, or surgery, with complete clinical data and confirmed pathological diagnosis. Exclusion criteria: (I) Lesions identified as mass-type breast lesions on ultrasound; (II) Patients with non-dense breast tissue as determined by ultrasound or mammography; (III) History of neoadjuvant chemotherapy, radiotherapy, or endocrine therapy for breast diseases prior to surgery; (IV) Unclear lesion nature or pathological diagnosis; (V) Poor image quality. This study was approved by the hospital's ethics committee of the Tenth Affiliated Hospital of Southern Medical University (Dongguan People's Hospital) (G-2023027) and adhered to the principles of the Declaration of Helsinki, thus informed consent from patients was waived.

Instruments and methods
A GE Logic E9 color Doppler ultrasound system with a high-frequency linear probe (L-16–5) was used, with a mechanical index of 0.16. Ultrasound examinations were performed by radiologists with over 10 years of experience. Patients were positioned supine with their arms raised at a 90° angle above their heads to fully expose the breasts and both axillae; if necessary, patients were asked to turn to their side. For all cases identified as non-mass lesions (NMLs) within a dense breast background via conventional ultrasound, 2D images were employed to assess lesion size, internal echo homogeneity, the presence of microcalcifications, posterior acoustic attenuation, and distortion of surrounding structures. Color Doppler imaging was utilized to evaluate the vascular characteristics of the lesions. The ultrasound image representing the lesion's longest axis was selected and exported in DICOM format, preserving full bit-depth in 16-bit grayscale, with a raw pixel value range of 0–4095, as configured in the Logic E9 system. To ensure reproducibility, all DICOM files retained their original acquisition parameters, including a dynamic range of 60 dB, time gain compensation (TGC) settings, and spatial compounding configurations. Prior to feature extraction, pixel values were linearly normalized to a [0, 1] range using min–max scaling, based on the device-specific maximum quantization units. Prior to feature extraction, pixel values were linearly normalized to a [0, 1] range using min–max scaling, based on the device-specific maximum quantization units[18,19].

Image data acquisition and processing
Using a single-blind method, ultrasound images were imported into Itk-Snap[20] (v4.20, http://www.itksnap.org/) by two radiologists with over 10 years of experience. The tumor boundaries were manually delineated along the edges, and the region of interest (ROI) was defined, which included spiculation, lobulation, and surrounding satellite or radial features. Using the ROI margin as the baseline, morphological adaptive segmentation was
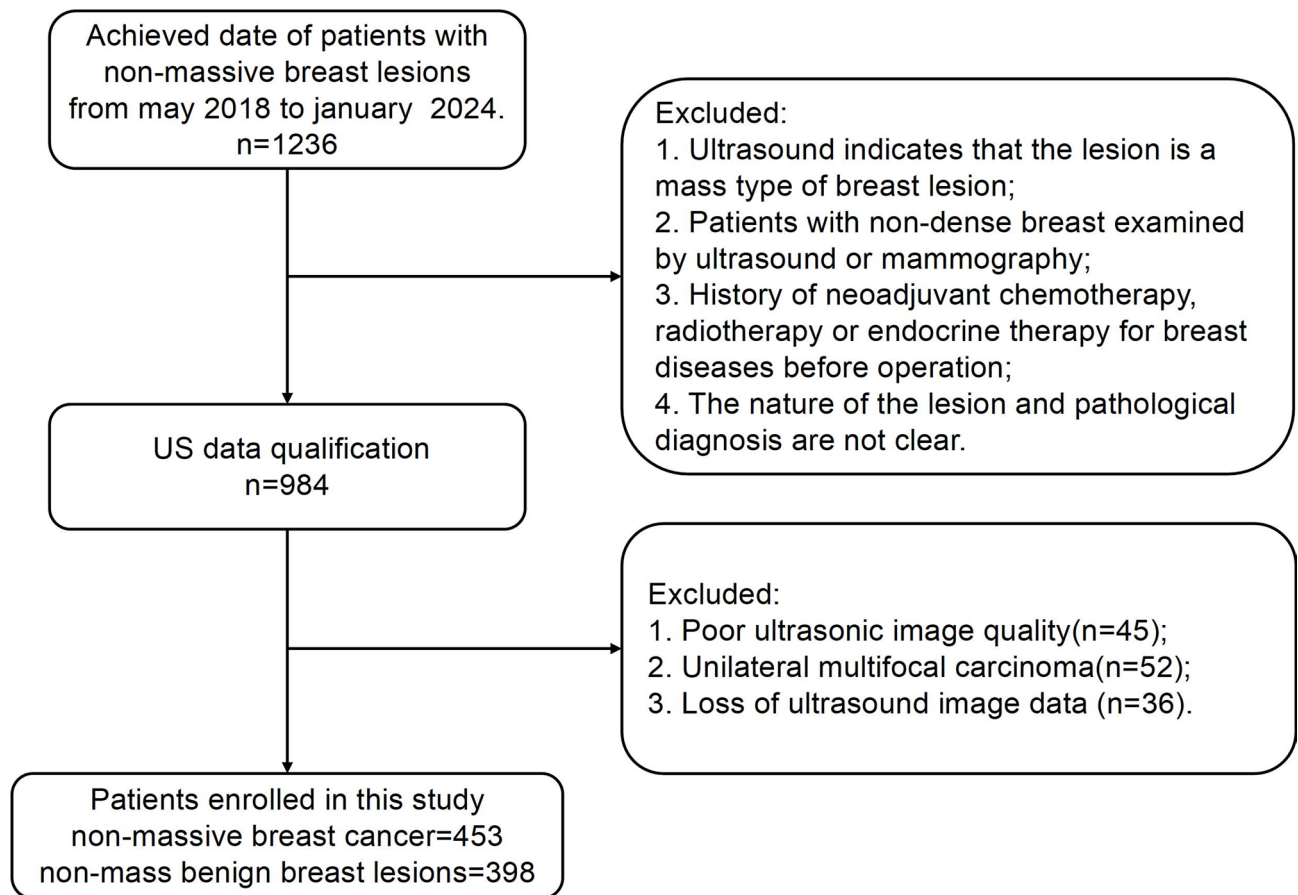
Fig. 1. Flowchart of patient enrollment.

applied to expand the peritumoral regions outward by 1 mm, 2 mm, 3 mm, 4 mm, and 5 mm. This yielded images of peri_1mm, peri_2mm, peri_3mm, peri_4mm, peri_5mm, and intratumoral regions (see Fig. 2). The interclass correlation coefficient (ICC) was used to assess the consistency of feature extraction between and within observers. ICC ratings were classified as follows: less than 0.40 was considered "Poor", 0.40–0.59 as "Fair", 0.60–0.74 as "Good", and 0.74–1.00 as "Excellent".

Construction of the radiomics model

The delineated ROI images were imported into pyradiomics[21] (v3.0.1, https://github.com/AIM-Harvard/pyradiomics), an open-source software, where image feature extraction was performed on the ultrasound images using a Wavelet filter while retaining the original images. This included morphological features, first-order features, and texture features, such as Gray Level Size Zone Matrix (GLSZM), Gray Level Co-occurrence Matrix (GLCM), Neighbouring Gray Tone Difference Matrix (NGTDM), Gray Level Dependence Matrix (GLDM), and Gray Level Run Length Matrix (GLRLM).

Extracted radiomic features were standardized using Z-score normalization. To mitigate potential multicollinearity among radiomic features, a t-test was conducted on features within the training set, retaining only those with a P-value < 0.05. Pearson correlation coefficients were computed to evaluate inter-feature relationships, and features with correlation coefficients ≥ 0.9 were considered redundant and excluded. To assess the necessity for multiple testing correction, a sensitivity analysis was additionally performed using the Bonferroni correction ($\alpha = 0.05/851$). Based on the minimum squared error (MSE) criterion, least absolute shrinkage and selection operator (LASSO) regression combined with ten-fold cross-validation was applied to select the optimal subset of radiomic features, further eliminating collinear variables by assigning zero-valued coefficients. Each selected feature was multiplied by its corresponding regression coefficient, and the resulting values were summed to generate a radiomics score for each patient.

Construction and evaluation of the predictive models

Based on the selected optimal radiomic features, thirteen algorithms, including Random Forest (RF), Decision Tree (DT), Extra Trees (ET), Support Vector Machine (SVM), Logistic Regression (LR), Stochastic Gradient Descent (SGD), K-Nearest Neighbors (KNN), XGBoost, AdaBoost, Gradient Boosting Decision Tree (GBDT), CatBoost, LightGBM, and Bayes, were used to construct intratumoral and peritumoral radiomics models. Ten-fold cross-validation was applied to identify the optimal parameters for each model, and Receiver Operating
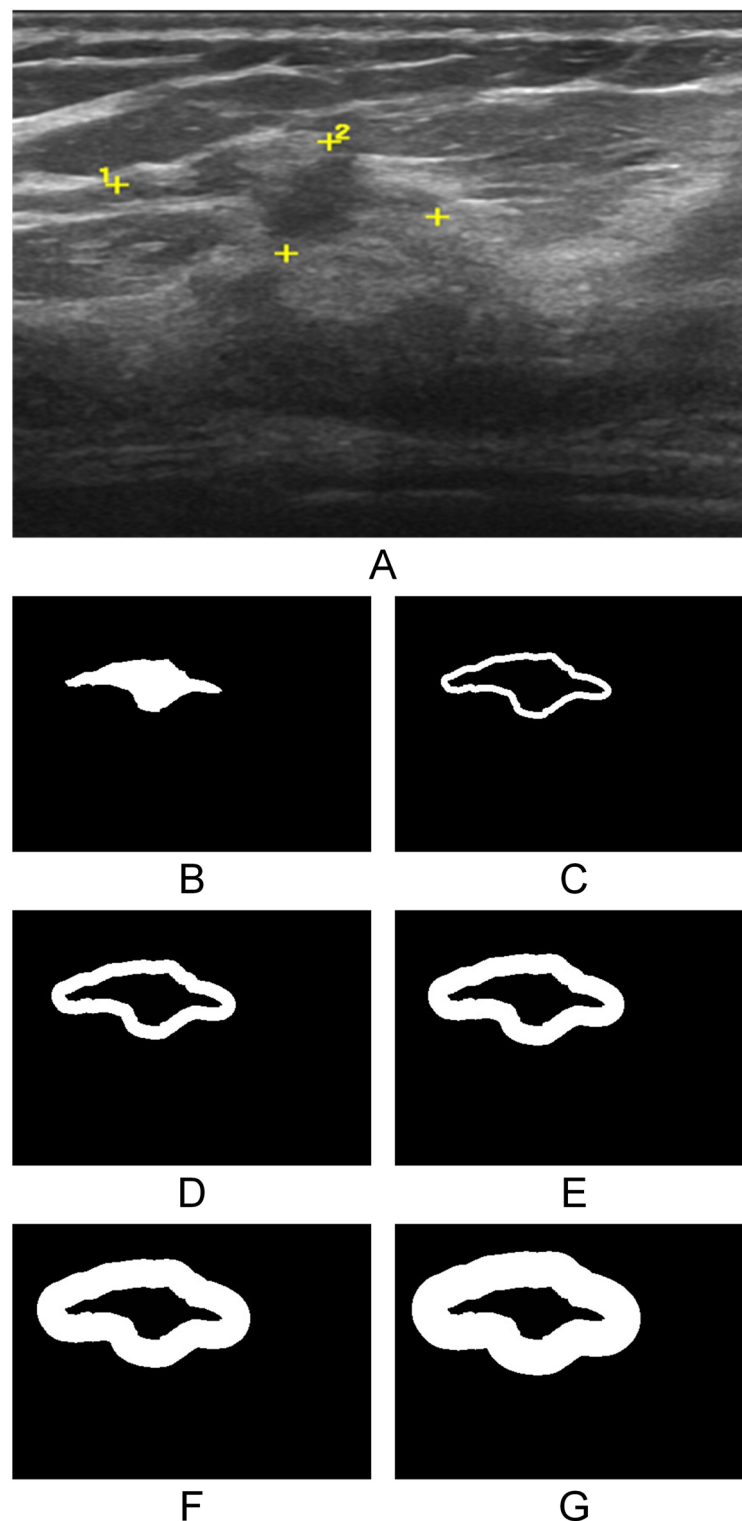
Fig. 2. Sketch of NML image. Ultrasound original image (**A**) with intratumoral mask (**B**), and peritumoral masks at 1 mm (**C**), 2 mm (**D**), 3 mm (**E**), 4 mm (**F**), and 5 mm (**G**).

Characteristic (ROC) curves were plotted to compare the Area Under the Curve (AUC) for all fourteen radiomics models.

To ensure that the performance of the multivariate model was not predominantly influenced by a single feature, univariate logistic regression was performed for each radiomic feature to evaluate its individual discriminative ability, using area under the curve (AUC) and P-value as metrics. Features with a univariate AUC greater than 0.70 were regarded as strong independent predictors. Additionally, DeLong's test was used to compare the AUC

of the optimal multivariate model with: (a) each top-performing univariate feature (AUC > 0.70) and (b) a reduced multivariate model excluding the top univariate feature. A statistically significant difference ($P < 0.05$) would indicate that the complete multivariate model offers superior predictive performance compared to any individual feature.

Univariate and multivariate logistic regression analyses were conducted using clinical and ultrasound features from the training group to identify independent predictors associated with NML, leading to the construction of clinical and ultrasound models. The radiomics model with the best predictive performance, combined with independent clinical and ultrasound predictors, was selected to construct a combined model. ROC curves were plotted, and the AUC, accuracy, sensitivity, and specificity of each model were calculated to evaluate their predictive efficacy for NML. Calibration curves and Decision Curve Analysis (DCA) were employed to assess the goodness-of-fit and clinical value of the models. Based on clinical and ultrasound independent predictors, as well as the optimal radiomics model, regression coefficients (β-values) and odds ratios (ORs) for each variable were calculated. A multivariate logistic regression analysis was then conducted to construct the nomogram.

Statistical methods

Statistical analyses were performed using R (version 4.2.0, https://www.R-project.org) and Python (version 3.10, https://www.python.org). Model construction and hyperparameter tuning were conducted in Python using the "scikit-learn" (version 1.4.1) and "XGBoost" (version 2.0.3) libraries. Hyperparameter optimization and cross-validation were implemented via "GridSearchCV" to minimize overfitting and improve model accuracy. Ten-fold cross-validation was conducted exclusively within the training cohort for feature selection, hyperparameter tuning, and model optimization. Final performance metrics were evaluated exclusively on the independent test cohort, which was not involved in any phase of model development.

XGBoost is an ensemble tree technique based on the loss produced by weak decision tree learners. XGBoost was trained as a baseline model, followed by training the final model with optimized hyperparameters. The hyperparameters of the XGBoost model were tuned using the GridSearchCV function in Scikit-learn, with ten-fold cross-validation performed. The selected hyperparameters for optimization included learning_rate, gamma, max_depth, subsample, min_child_weight, and n_estimators. The GridSearchCV method in Scikit-learn with ten-fold cross-validation was also used to tune the hyperparameters of RF, DT, ET, SVM, LR, SGD, KNN, XGBoost, AdaBoost, GBDT, CatBoost, LightGBM, and Bayes.

Categorical data are presented as percentages (%) and comparisons were performed using the chi-square (χ^2) test. Continuous data were evaluated for normality using the Kolmogorov–Smirnov test. Normally distributed data are presented as mean ± standard deviation ($x \pm s$) and compared using the Student's t-test. Pearson correlation analysis was used to select relevant radiomic features. Receiver operating characteristic (ROC) curves were used to assess the predictive performance of each model, and the DeLong'test was applied to compare their relative performances. A nomogram prediction model was constructed, with the optimal threshold determined through ROC curve analysis. The threshold was defined by maximizing the Youden index. Based on the optimal cutoff values, the area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were calculated for each model. Calibration curves were employed to evaluate model consistency, and decision curve analysis (DCA) quantified the net benefits at various threshold probabilities to assess the clinical significance of the models. The significance level was set at α = 0.05 (two-tailed).

Results

Patient demographics

A total of 851 cases of breast NWL meeting the inclusion and exclusion criteria were identified from May 30, 2018, to January 20, 2024, with postoperative pathological results showing 453 malignant cases and 398 benign cases. Among these, the training group comprised 327 malignant and 269 benign patients, while the validation group included 126 malignant and 129 benign patients. The chi-square statistic was 1.920, $P = 0.166$, indicating no significant statistical difference. There were no significant differences in clinical and ultrasound characteristics between the two groups ($P > 0.05$; Supplementary Table S1). Detailed clinical and ultrasound characteristics of the patients are presented in Table 1. Results of the multivariate analysis are detailed in Supplementary Tables S2 and S3. There were significant correlations between patient age, weight, lesion location, lesion longest diameter, posterior echo attenuation, surrounding structural disruption, and blood flow characteristics with non-mass type breast cancer ($P < 0.05$). No significant differences were observed for other clinical and ultrasound characteristics ($P > 0.05$).

The above clinical and ultrasound independent predictive factors were incorporated into logistic regression to establish the clinical model and ultrasound model, which were evaluated using ROC curves (see Supplementary Fig. S1 online). The accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of these models are detailed in Supplementary Table S4.

Extraction, selection, and construction of radiomics features

A total of 851 radiomics features were extracted from the ultrasound images of patients with breast NWL using radiomics technology. Feature selection was performed using student's t-test, Pearson correlation test, and LASSO (see Supplementary Fig. S2 online). Correlation heatmaps of the peritumoral (1–5 mm) and intratumoral regions (see Supplementary Fig. S3 online) revealed distinct collinearity patterns. The results showed that after feature selection, the radiomic features exhibited low correlation ($|r| < 0.6$), confirming the minimization of residual collinearity. Additionally, Bonferroni correction reduced the number of significant features in each region by an average of 8.1%. Importantly, LASSO regression applied to both corrected and uncorrected feature pools yielded the same final predictive factors (Supplementary Table S5). This suggests that multivariate regularization inherently resolves the potential false-positive issues in univariate selection, thus negating the need for additional

	Train group				Validation group			
Variables	Benign NML (n = 269)	Malignant NML (n = 327)	$t/W/\chi^2$	P	Benign NML (n = 129)	Malignant NML (n = 126)	$t/W/\chi^2$	P
Age, Median (Q1, Q3)	40 (34, 49)	48 (42.5, 54)	26840[b]	< 0.001	41 (34, 49)	47 (39, 53)	5874.5[b]	< 0.001
High, Median (Q1, Q3)	160 (158, 165)	160 (158, 165)	45174[b]	0.567	160 (158, 165)	160 (158, 164)	8635.5[b]	0.386
Weigh, Median (Q1,Q3)	62 (57, 70)	60 (56, 67.67)	49,475.5[b]	0.009	63 (57, 70)	61.01 (56, 68)	9440.5[b]	0.026
BMI, Median (Q1, Q3)	1.66 (1.59, 1.74)	1.64 (1.57, 1.72)	49,133.5[b]	0.014	1.682 ± 0.126	1.644 ± 0.124	2.419[b]	0.016
Menarche age, Median (Q1, Q3)	12 (12, 13)	12 (12, 13)	41930[b]	0.248	12 (12, 13)	12 (12, 12.75)	8448.5[b]	0.492
Menopause, n (%)			26.542[c]	< 0.001			9.547[c]	0.002
No	234 (86.989)	225 (68.807)			114 (88.372)	91 (72.222)		
Yes	35 (13.011)	102 (31.193)			15 (11.628)	35 (27.778)		
Position, n (%)			12.06[c]	0.007			10.077[c]	0.018
Outer upper quadrant	142 (52.788)	210 (64.220)			70 (54.264)	84 (66.667)		
Inner upper quadrant	63 (23.420)	61 (18.654)			25 (19.380)	28 (22.222)		
Outer lower quadrant	41 (15.242)	45 (13.761)			21 (16.279)	10 (7.937)		
Inner lower quadrant	23 (8.550)	11 (3.364)			13 (10.078)	4 (3.175)		
Long diameter, Median (Q1, Q3)	2 (1.1, 4.36)	2.54 (1.7, 3.8)	38,264.5[b]	0.006	2.2 (1.4, 3.84)	2.8 (1.83, 3.8)	6991.5[b]	0.054
Wide diameter, Median (Q1, Q3)	0.9 (0.59, 1.6)	1.1 (0.8, 1.6)	36525[b]	< 0.001	1 (0.6, 1.7)	1.2 (0.9, 1.6)	6654.5[b]	0.012
Focus orientation, n (%)			0.020[c]	0.887			Fisher	1
Parallel	261 (97.026)	319 (97.554)			126 (97.674)	123 (97.619)		
No parallel	8 (2.974)	8 (2.446)			3 (2.326)	3 (2.381)		
Rear echo, n (%)			Fisher	< 0.001			Fisher	< 0.001
No change	237 (88.104)	256 (78.287)			120 (93.023)	98 (77.778)		
Attenuation	27 (10.037)	71 (21.713)			7 (5.426)	28 (22.222)		
Enhancement	5 (1.859)	0 (0)			2 (1.550)	0 (0)		
Calcification , n (%)			90.702[c]	< 0.001			45.974[c]	< 0.001
No calcification	211 (78.439)	150 (45.872)			86 (66.667)	48 (38.095)		
Coarse calcification	28 (10.409)	22 (6.728)			24 (18.604)	9 (7.143)		
Fine calcification	30 (11.152)	155 (47.401)			19 (14.729)	69 (54.762)		
Surrounding structure, n (%)			21.947[c]	< 0.001			20.787[c]	< 0.001
No twist	147 (54.647)	115 (35.168)			71 (55.039)	33 (26.190)		
Twist	122 (45.353)	212 (64.832)			58 (44.961)	93 (73.810)		
Blood flow characteristics, n (%)			48.509[c]	< 0.001			21.372[c]	< 0.001
Grade 0	119 (44.238)	75 (22.936)			56 (43.411)	28 (22.222)		
Grade I	76 (28.253)	73 (22.324)			33 (25.581)	24 (19.048)		
Grade II	40 (14.870)	93 (28.440)			24 (18.605)	39 (30.952)		
Grade III	34 (12.639)	86 (26.300)			16 (12.403)	35 (27.778)		
Peripheral blood flow, n (%)			2.636[c]	0.104			Fisher	0.066
None	257 (95.539)	321 (98.165)			122 (94.574)	125 (99.206)		
Yes	12 (4.461)	6 (1.835)			7 (5.426)	1 (0.794)		

Table 1. Comparison of clinical and ultrasound data of benign and malignant patients with NML between training group and validation group. a: *t*-value, b: *W*-value , c: χ^2-value, NML: non-mass breast lesions.

corrections. The optimal radiomics features were obtained from the peri-lesional (1 mm, 2 mm, 3 mm, 4 mm, 5 mm) and intra-lesional models with corresponding Lambda values of 0.009, 0.004, 0.002, 0.002, 0.001, and 0.003, respectively. The results are detailed in Supplementary Table S6.

Thirteen algorithms, including RF, DT, ET, SVM, LR, SGD, KNN, XGBoost, Adaboost, GBDT, CatBoost, LightGBM, and Bayes, were used to construct the ultrasound radiomics models. The results are detailed in Table 2 and Supplementary Tables S7–S11. The predictive task aimed to differentiate between benign and malignant non-mass breast lesions (NML). As shown in the ROC curves (see Supplementary Fig. S4 online), the intratumoral model (RS_in_XGBoost) achieved an AUC of 0.842 (95% CI 0.788–0.896), while peri-lesional models exhibited varying performance: the 1 mm peri-lesional model (RS_peri_1mm_XGBoost) yielded an AUC of 0.819 (95% CI 0.762–0.876), followed by the 2 mm peri-lesional model (RS_peri_2mm_RF) with an AUC of 0.799 (95% CI 0.739–0.858). Notably, higher performance was observed in the 3 mm (RS_peri_3mm_XGBoost: AUC = 0.825, 95% CI 0.769–0.882), 4 mm (RS_peri_4mm_SVM: AUC = 0.904, 95% CI 0.860–0.948), and 5 mm (RS_peri_5mm_LR: AUC = 0.876, 95% CI 0.828–0.926) peri-lesional regions, with the 4 mm SVM-based model demonstrating the highest discriminative capability.

Radiomics models[a]	AUC (95%CI)	ACC (95%CI)	SEN (95%CI)	SPE (95%CI)	PPV (95%CI)	NPV (95%CI)
RS(in_RF)	0.793 (0.733–0.853)	0.736 (0.671–0.801)	0.682 (0.613–0.751)	0.789 (0.728–0.850)	0.759 (0.696–0.823)	0.717 (0.650–0.784)
RS(in_DT)	0.619 (0.508–0.731)	0.645 (0.535–0.755)	0.976 (0.936–1.015)	0.257 (0.157–0.358)	0.606 (0.494–0.718)	0.900 (0.830–0.970)
RS(in_ET)	0.778 (0.717–0.840)	0.730 (0.664–0.796)	0.727 (0.661–0.793)	0.733 (0.668–0.799)	0.727 (0.661–0.793)	0.733 (0.668–0.799)
RS(in_SVM)	0.791 (0.730–0.851)	0.753 (0.689–0.817)	0.670 (0.601–0.740)	0.833 (0.778–0.889)	0.797 (0.738–0.857)	0.722 (0.655–0.788)
RS(in_LR)	0.631 (0.559–0.702)	0.626 (0.554–0.697)	0.748 (0.683–0.812)	0.461 (0.387–0.534)	0.653 (0.582–0.723)	0.574 (0.501–0.647)
RS(in_SGD)	0.716 (0.649–0.783)	0.682 (0.613–0.750)	0.659 (0.589–0.729)	0.703 (0.636–0.771)	0.682 (0.614–0.751)	0.681 (0.612–0.750)
RS(in_KNN)	0.733 (0.668–0.799)	0.682 (0.613–0.750)	0.652 (0.582–0.723)	0.713 (0.646–0.780)	0.706 (0.638–0.773)	0.660 (0.590–0.760)
RS(in_XGBoost)	0.805 (0.746–0.863)	0.764 (0.701–0.827)	0.716 (0.649–0.783)	0.811 (0.753–0.869)	0.786 (0.727–0.848)	0.745 (0.680–0.810)
RS(in_Adaboost)	0.628 (0.557–0.700)	0.615 (0.543–0.686)	0.871 (0.821–0.921)	0.337 (0.267–0.407)	0.587 (0.514–0.660)	0.707 (0.640–0.775)
RS(in_GBDT)	0.778 (0.717–0.840)	0.758 (0.695–0.822)	0.682 (0.613–0.751)	0.833 (0.778–0.889)	0.800 (0.741–0.859)	0.728 (0.662–0.794)
RS(in_CatBoost)	0.749 (0.685–0.814)	0.682 (0.613–0.750)	0.522 (0.448–0.596)	0.851 (0.798–0.904)	0.787 (0.726–0.848)	0.627 (0.556–0.699)
RS(in_LightGBM)	0.716 (0.649–0.783)	0.659 (0.589–0.729)	0.556 (0.482–0.629)	0.817 (0.760–0.87)	0.822 (0.765–0.879)	0.547 (0.474–0.621)
RS(in_Bayes)	0.708 (0.640–0.775)	0.721 (0.654–0.787)	0.912 (0.869–0.954)	0.468 (0.394–0.541)	0.694 (0.626–0.762)	0.800 (0.741–0.859)

Table 2. Performance of intratumoral radiomics models in the validation group. [a]The intratumoral radiomics model was built using 13 algorithms: Random Forest (RF), Decision Tree (DT), Extra Trees (ET), Support Vector Machine (SVM), Logistic Regression (LR), Stochastic Gradient Descent (SGD), K-Nearest Neighbors (KNN), XGBoost, AdaBoost, Gradient Boosting Decision Tree (GBDT), CatBoost, LightGBM, and Bayes. AUC: area under curve, ACC: accuracy, SEN: sensitivity ,SPE: specificity, PPV: positive predictive Value, NPV: negative predictive value, CI: confidence interval.

Univariate analysis and multivariate model robustness

Univariate logistic regression identified 26, 36, 60, 73, 96, and 45 radiomic features with significant discriminative power in the peritumoral (1–5 mm) and intratumoral regions ($P < 0.05$; Supplementary Tables S12–S17). The DeLong'test indicated that the optimal multivariate model for each region significantly outperformed all the best-performing univariate features ($P < 0.05$), confirming that no single feature dominated the model.

Feature analysis of radiomics model

Seven "intersection features" were identified from the overlap of 26, 36, 60, 73, 96, and 45 imaging features extracted from the ultrasound images of the aforementioned different ROI regions: "wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis", "original_glcm_JointAverage", "original_firstorder_Minimum", "wavelet-LLH_firstorder_Kurtosis", "wavelet-HLH_glrlm_RunVariance", "wavelet-LHH_glrlm_RunVariance", and "wavelet-LLL_gldm_LargeDependenceHighGrayLevelEmphasis". Among these features, the effects observed in the intratumoral and peritumoral 4 mm imaging models were oppositional, particularly for "wavelet-HLH_glrlm_RunVariance" and "wavelet-LHH_glrlm_RunVariance". In the intratumoral model, smaller values of these two features were associated with a higher probability of predicting non-massive breast cancer, whereas in the peritumoral 2 mm model, larger values were linked to a higher probability of the same prediction. While "wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis" and "wavelet-LLL_gldm_LargeDependenceHighGrayLevelEmphasis" have a minimal impact on the intratumoral model predictions, in the peritumoral 2 mm model, a smaller value of "wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis" and a larger value of "wavelet-LLL_gldm_LargeDependenceHighGrayLevelEmphasis" significantly influence the predictions. While "wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis" and "wavelet-LLL_gldm_LargeDependenceHighGrayLevelEmphasis" had minimal impact on the intratumoral model predictions, in the peritumoral 2 mm model, a smaller value of "wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis" and a larger value of "wavelet-LLL_gldm_LargeDependenceHighGrayLevelEmphasis" significantly influenced the predictions.

Figure 3 visualizes the feature selection results of the optimal radiomics models for the six ROI groups, where the six red nodes represent the selected features for each group's model. Notably, the seven light pink nodes were consistently selected by the best radiomics models across all six ROI groups, indicating their robustness and potential significance in model predictive performance. In contrast, features such as "wavelet-LLH_glszm_SizeZoneNonUniformity" were only selected by the peri-lesional 5 mm radiomics model, suggesting that their importance may be specific to that model. A partial dependence plot was created for the aforementioned seven "intersection features" based on the intra-lesional and peri-lesional 4 mm radiomics models.

Nonlinear spatial dynamics of radiomics signatures

The coefficient magnitude (β) displayed a triphasic pattern across peritumoral regions: a sharp decline from intratumoral ($\beta = 0.053$) to peri_1mm ($\beta = 0.029$, $\Delta = 45.3\%$), followed by further decay in peri_2mm ($\beta = 0.011$, $\Delta = 62.1\%$), and stabilization beyond peri_3mm ($\beta \approx 0.02$). However, the gradient decay index (GDI) exhibited an inverse trend: the maximal signal shift occurred between peri_1mm and peri_2mm (GDI = 0.818), while peri_3mm and beyond showed no measurable decay (GDI = 0) (Supplementary Table S18). This paradox suggests distinct biological regimes: early-stage tumor-stroma interactions dominate the peri_1-2 mm region (high GDI), while more distant regions (peri_3-5 mm) reach a steady-state microenvironment, homogenized by systemic immune regulation or mechanical constraints.
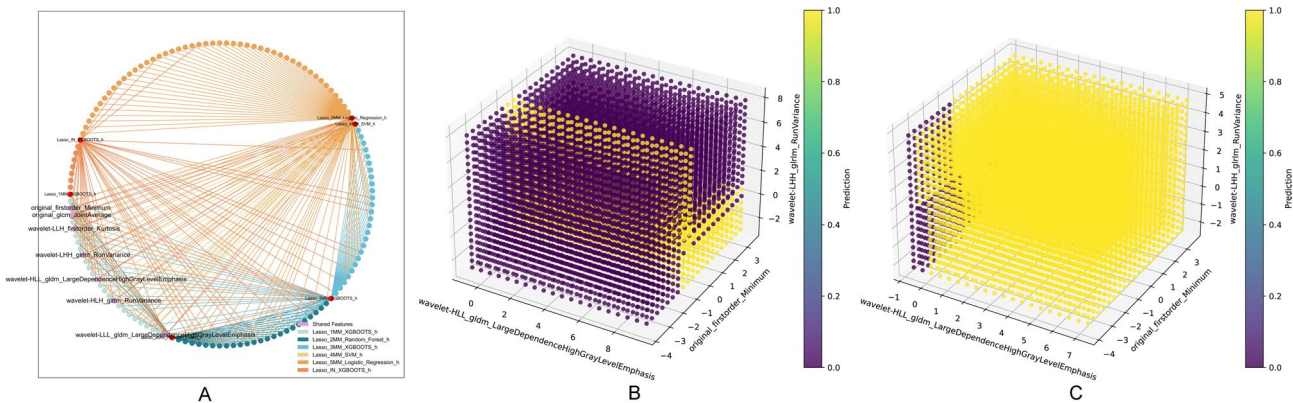
Fig. 3. Feature analysis of Radiomics model. (**A**) Light-colored nodes indicate the best radiomics features for the respective ROIs, while purple nodes represent features shared across multiple models. The edges color-coded by model indicate which features were selected by which models. (**B**) and (**C**) illustrate the partial dependence of the model on three features (wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis, original_firstorder_Minimum, wavelet-LHH_glrlm_RunVariance). Each axis represents the range of a feature value, and the color of the points reflects the model's predictions. The color gradient allows for a visual assessment of the combined effect of these three features on the predictions.

Combined models[a]	AUC (95% CI)	ACC (95% CI)	SEN (95% CI)	SPE (95% CI)	PPV (95% CI)	NPV (95% CI)	AUDC (95%CI)
All (peri_1mm_XGBoost)	0.799 (0.707–0.891)	0.740 (0.640–0.840)	0.889 (0.816–0.962)	0.610 (0.499–0.721)	0.667 (0.559–0.774)	0.862 (0.782–0.942)	0.104 (−0.274–0.251)
All (peri_2mm_RF)	0.850 (0.767–0.932)	0.805 (0.714–0.896)	0.882 (0.808–0.957)	0.744 (0.644–0.844)	0.732 (0.630–0.833)	0.889 (0.816–0.962)	0.294 (−0.530–0.549)
All (peri_3mm_XGBoost)	0.833 (0.748–0.919)	0.753 (0.655–0.852)	0.723 (0.621–0.826)	0.800 (0.708–0.892)	0.850 (0.768–0.932)	0.649 (0.540–0.757)	0.160 (−0.354–0.361)
All (peri_4mm_SVM)	0.782 (0.688–0.877)	0.740 (0.640–0.840)	0.771 (0.675–0.868)	0.714 (0.611–0.817)	0.692 (0.587–0.798)	0.789 (0.696–0.883)	0.160 (−0.292–0.262)
All (peri_5mm_LR)	0.830 (0.744–0.916)	0.805 (0.714–0.896)	0.769 (0.673–0.865)	0.842 (0.758–0.926)	0.833 (0.748–0.919)	0.780 (0.686–0.875)	0.205 (−0.323–0.294)
All (in_XGBoost)	0.858 (0.778–0.939)	0.792 (0.699–0.885)	0.756 (0.657–0.854)	0.844 (0.760–0.927)	0.872 (0.794–0.949)	0.711 (0.607–0.814)	0.282 (−0.306–0.297)

Table 3. Performance of combined models in the validation group. [a]The optimal radiomics models from each ROI group were combined with independent clinical and ultrasound predictive factors. AUC: area under curve, ACC: accuracy, SEN: sensitivity ,SPE: specificity ,PPV: positive predictive Value, NPV: negative predictive value, AUDC: area under the decision curve analysis, CI: confidence interval, RF: random forest, SVM: Support Vector Machine, LR: Logistic Regression.

Effectiveness evaluation of combined model

All model comparisons are detailed in Supplementary Table S19, showing that the combined model outperformed other individual models ($P < 0.01$), with accuracy, sensitivity, specificity, positive predictive value, and negative predictive value detailed in Table 3.

The ROC curves for the six combined models show that the joint model constructed with intra-lesional radiomics, clinical, and ultrasound independent predictors had the best predictive performance, with an AUC of 0.858 (95% CI 0.778–0.939). The next best was the combined model using peri-lesional 2 mm radiomics, with an AUC of 0.850 (0.767–0.932). The calibration curve indicates that this model has good fit in predicting TNBC ($P > 0.05$). The DCA curve shows that the combined model using the peri-lesional 2 mm radiomics and clinical, ultrasound independent predictors had higher clinical net benefit within the risk threshold range of 0.069 to 0.966, with an Area Under the Decision Curve Analysis (AUDC) of 0.294 (95%: −0.530–0.549). The second-best was the combined model with intra-lesional radiomics, which had an AUDC of 0.282 (95%: −0.306–0.297) (see Fig. 4).

The nomogram constructed using the intra-lesional radiomics model and clinical, ultrasound independent predictors had a diagnostic performance C-index of 0.982. When the nomogram score threshold is 167.808, the optimal classification prediction model can be obtained. A score greater than 167.808 predicts non-massive breast cancer, while a score less than 167.808 predicts non-mass benign breast lesions. The nomogram constructed using the peri-lesional 2 mm radiomics model and clinical, ultrasound independent predictors had a diagnostic performance C-index of 0.978. When the nomogram score threshold is 110.745, the optimal classification prediction model can be obtained. A score greater than 110.745 predicts non-massive breast cancer, while a score less than 110.745 predicts non-mass benign breast lesions (see Fig. 5).
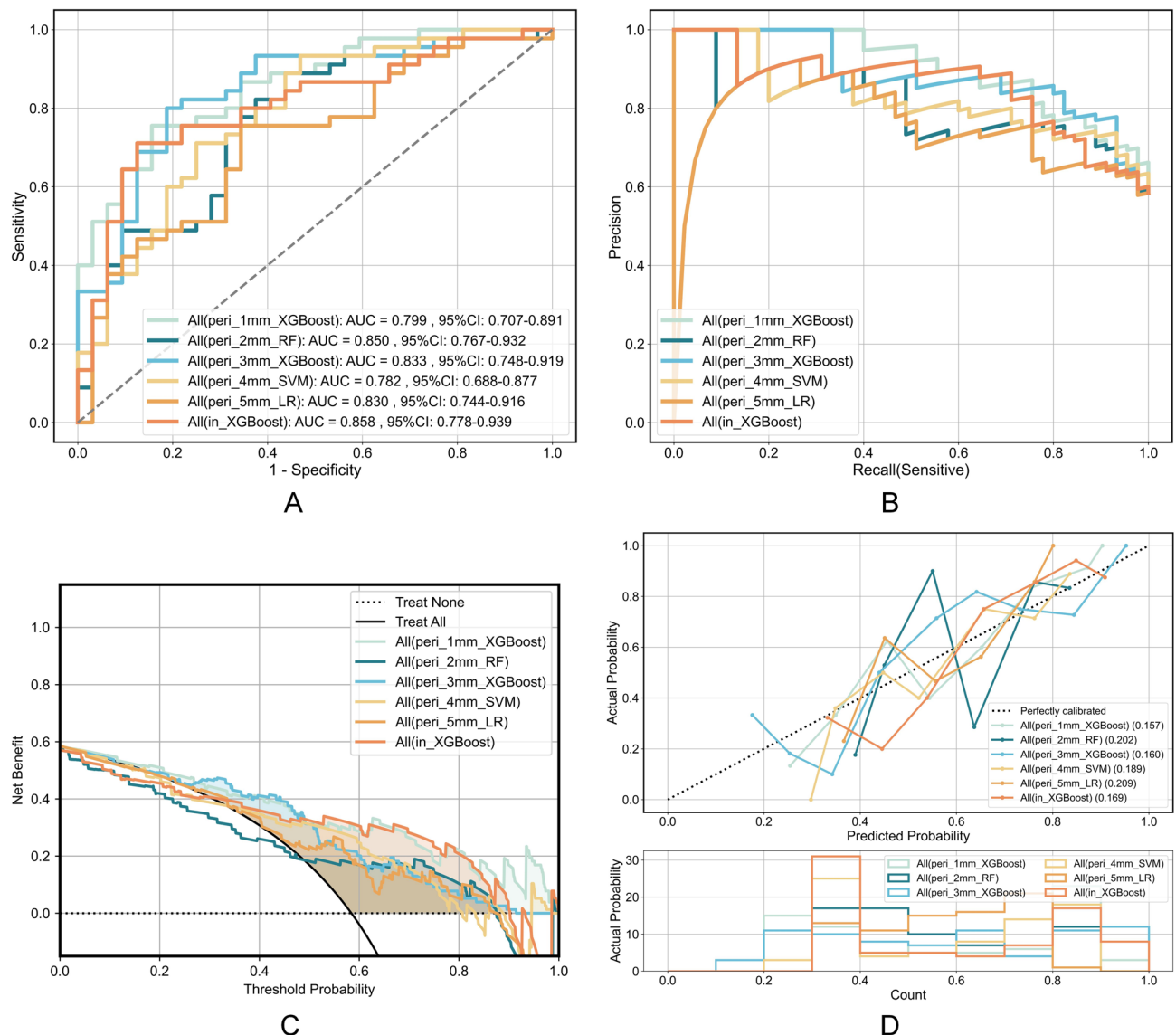
Fig. 4. Performance of combined models in the validation group. The optimal radiomic models for the validation group, including the intratumoral and peritumoral areas at 1 mm, 2 mm, 3 mm, 4 mm, and 5 mm, were combined with independent clinical and ultrasound predictors to obtain the ROC curve (**A**), P–R curve (**B**), DCA curve (**C**), and calibration curve (**D**). ROC: receiver operating characteristic, AUC: area under curve, CI: confidence interval, P-R: precision- recall, DCA: decision curve analysis.

Models display and application

To facilitate the application of our findings for clinicians, researchers, patients, and their families, we have developed this diagnostic prediction system, which can be accessed at the following website: https://jarrily9527 nml.streamlit.app/.

Discussion

Compared to breast mass lesions, the detection and differentiation of benign and malignant NWL are more challenging. According to recent literature statistics[22,23], benign NWL accounts for 53.8%, while malignant NWL constitutes 46.2%. There is an overlap in the ultrasound features of NWL, with structural distortion and ductal changes being common characteristics[24–26]. Especially in the context of dense breast tissue, the denser the fibroglandular tissue, the more the lesions overlap with it, making the differentiation of benign and malignant NWL even more difficult.

In this study, age is identified as a risk factor for malignant NWL. Previous research has indicated that advancing age is often associated with changes in metabolism and immune function, the accumulation of oncogenes, and insufficient DNA methylation, all of which are closely related to the occurrence of breast cancer[27]. The length of the lesion, the presence of microcalcifications, structural distortion, and blood flow characteristics are independent predictive factors for malignant lesions. Additionally, blood flow characteristics are an independent
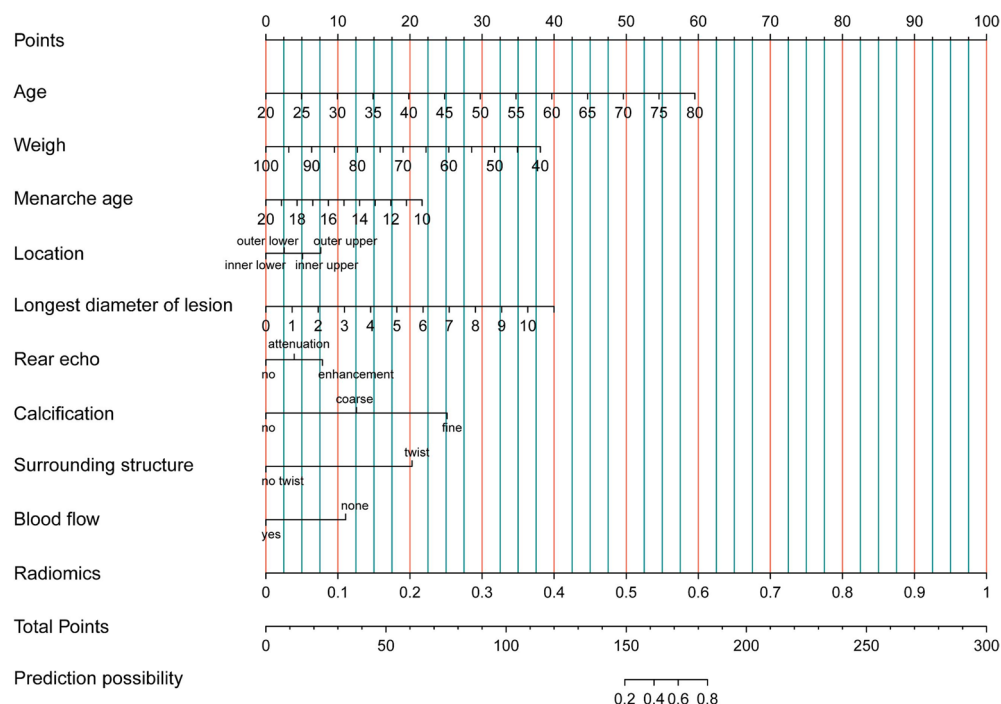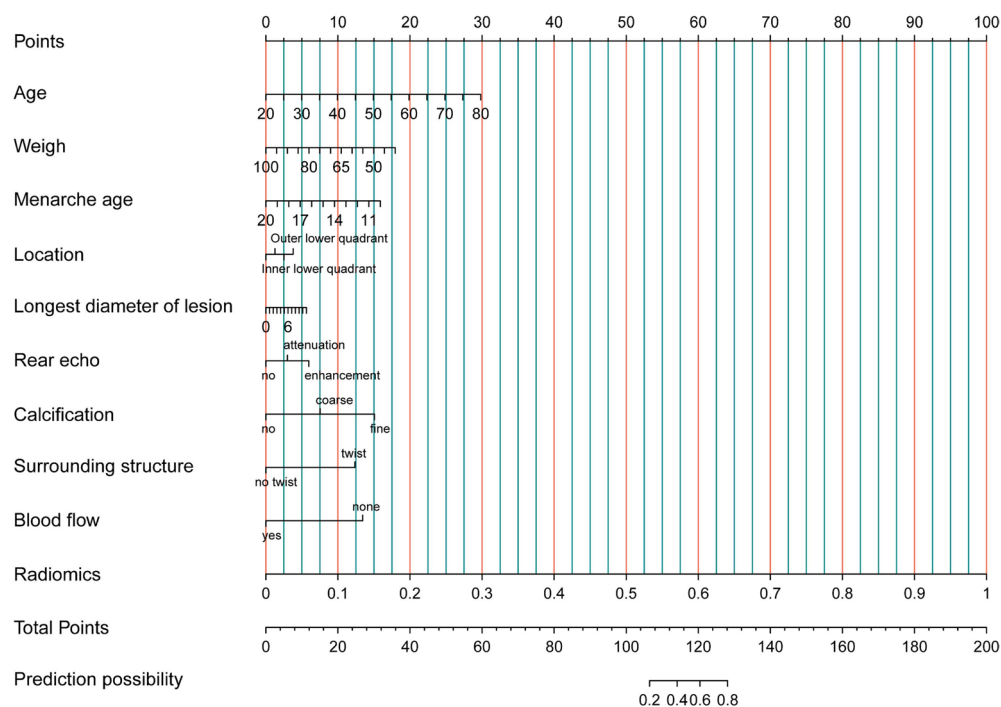
Fig. 5. Nomograms were established based on clinical and ultrasound independent predictors combined with the intratumoral (**A**) and peritumoral 2 mm (**B**) radiomic models in the validation group. The predictors included clinical and ultrasound independent factors as well as the radiomics score. A vertical line was drawn from each predictor to the "score" to obtain the score for each predictor. The scores for each predictor were then summed. The "total score" corresponds to the probability of the malignancy of NML.

risk factor; malignant lesions typically exhibit blood flow signals rated at 2 to 3, whereas benign lesions more frequently present signals rated at 0 to 1. Studies have shown that during the invasive growth of breast tumors, tumor cells grow rapidly, and the marginal area experiences hypoxia, which promotes the growth of microvessels within the lesion and the formation of a microvascular network[28]. This study found that structural distortion in the breast is also considered an important factor associated with malignancy. However, this differs from previous research[29], possibly due to differences in the pathological types of the study subjects. While structural distortion and ductal changes are common features of both benign and malignant NWL, structural distortion is more prevalent in the ultrasound images of malignant lesions compared to benign ones[30].

High-frequency breast ultrasound has been widely used in the screening and diagnosis of breast cancer, significantly increasing the detection rate of breast cancer in dense breast tissue[31]. However, the ultrasound model in this study demonstrated low diagnostic performance, confirming the poor ultrasound diagnostic capability for NML in dense breast tissue. The detection rate is low due to the ultrasound features of NWL differing from those of typical mass-type breast cancer. Previous studies have described NML as regions of diffuse echotexture changes that differ from the surrounding breast tissue[25]. In dense breast tissue, the denser the fibroglandular tissue, the more the lesions overlap with it, which increases the difficulty of detecting non-mass breast cancer. Ultrasound examination of superficial small organs such as the breast primarily relies on the operator's technique and experience. As a result, the differentiation of benign and malignant breast NWL is influenced by subjective factors such as the reporting physician's experience. There is a need for an objective tool to improve the detection rate of non-mass breast cancer in dense breast tissue.

In recent years, imaging omics methods have been used to explore the relationship between the peritumoral imaging features of tumors and their surrounding microenvironments. Several imaging biomarkers reflecting information about the tumor microenvironment have been identified, demonstrating significant value in tumor differential diagnosis, metastasis prediction, treatment evaluation, and prognosis prediction. Sun et al.[32] found that in ultrasound images of breast cancer, a lymph node metastasis prediction model combining intratumoral and peritumoral imaging features significantly outperformed models based solely on intratumoral or peritumoral features. Moreover, the model constructed using CNN achieved the best performance, clearly surpassing other models. Guo et al.[33] attempted to use dual-modal images from conventional breast ultrasound and strain elastography to conduct radiomics research on benign and malignant breast lesions. They extracted imaging features from the intratumoral, peritumoral 5 mm, and "intratumoral + peritumoral" regions to construct predictive models for breast lesions. The results indicated that the model combining grayscale ultrasound, strain elastography dual-modal images, and "intratumoral + peritumoral" imaging features had the highest diagnostic performance. Furthermore, scholars such as Huang et al.[34] combined photoacoustic imaging with intratumoral and peritumoral radiomics techniques to determine that a 5 mm peritumoral region is the optimal area, revealing its critical role in enhancing the predictive accuracy of malignant breast nodules when combined with intratumoral features. In a multimodal study, Liu et al.[35] demonstrated that the peritumoral radiomics model (GPTV_radiomics) and deep learning model (GPTV_DL) based on multiparametric MRI showed significantly superior performance in predicting lymphovascular invasion in the peritumoral region compared to intratumoral analysis (AUC: 0.771 vs. 0.720). An MRI study by Yu et al.[36] further validated this finding, with their peritumoral features predicting lymphovascular invasion reaching an AUC of 0.83, an 8% improvement over intratumoral features. These studies collectively reveal that the peritumoral microenvironment at 4-5 mm may influence tumor invasiveness through mechanisms such as angiogenesis and immune infiltration. Its texture features demonstrate diagnostic advantages across multimodal imaging platforms, including ultrasound, photoacoustic, and MRI. Particularly in lesions with indistinct borders, peritumoral heterogeneity may more sensitively reflect tumor biological activity, providing a solid theoretical foundation for constructing peritumoral models.

This study constructed 13 radiomics models based on intratumoral and peritumoral ultrasound images at depths of 1 mm, 2 mm, 3 mm, 4 mm, and 5 mm, utilizing high-throughput radiomics technology. Among these, the peritumoral 4 mm radiomics model constructed using the SVM algorithm exhibited the best predictive performance, with an AUC of 0.904 (95% CI: 0.860–0.948), while the intratumoral radiomics model constructed using the XGBoost algorithm had an AUC of 0.842 (95% CI: 0.788–0.896). The DeLong'test for the two models showed a P-value of 0.773, indicating no statistically significant difference. However, the results from the ROC curves reveal that the hidden biological information carried by the peritumoral 4 mm region is more valuable than that of the intratumoral and other peritumoral regions (AUC_peri_4mm = 0.904 > AUC_inter = 0.842).

The effectiveness of "wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis" in the peritumoral 2 mm model is greater than that of "original_glcm_JointAverage" and "original_firstorder_Minimum," a trend not observed in the intratumoral model. Several factors contribute to these results: the observed spatial heterogeneity of radiomics features likely arises from intrinsic differences in tumor biology and imaging characteristics. In intratumoral regions, densely packed malignant cells and necrotic zones may result in homogeneous texture patterns, which explains why lower wavelet-HLH_glrlm_RunVariance values (reflecting reduced pixel variation) are associated with malignancy. In contrast, peritumoral regions (1-2 mm) often show infiltrative tumor growth and stromal reactions, increasing tissue complexity. Higher RunVariance values in these regions may capture disrupted architectural patterns caused by tumor invasion, aligning with their association with aggressive phenotypes. Furthermore, features such as wavelet-HLL_gldm_LargeDependenceHighGrayLevelEmphasis, which quantify the clustering of high-intensity pixels, may reflect microvascular proliferation in the peritumoral stroma—a hallmark of tumor angiogenesis. This is supported by previous studies linking peritumoral vascularity to malignancy risk[37,38]. The reduced role of these features intratumorally may be attributed to necrotic dominance, which suppresses gray-level dependency[39]. Collectively, these findings support the "seed and soil" hypothesis, wherein the peritumoral microenvironment (2-4 mm) plays a crucial role in determining tumor behavior, and ultrasound radiomics offers a non-invasive method to decode this spatial interplay.

The observed β and GDI patterns challenge the classical monotonic decay hypothesis. We propose that tumor-secreted immunosuppressors lead to homogenization of the proximal zones, while the peri_1-2 mm region functions as a diagnostic 'twilight zone,' where increasing stromal heterogeneity induces abrupt signal shifts, as captured by wavelet features such as wavelet-HLH_glrlm_RunVariance. Beyond the peri_3mm region, systemic homeostasis (e.g., collagen crosslinking, immune surveillance) mitigates variability, rendering extended sampling unnecessary.

Among the combined models constructed using intratumoral and various peritumoral ultrasound imaging features along with clinical and ultrasound independent predictive factors, the combined model with the peritumoral 2 mm region achieved the highest AUC, improving from 0.799 to 0.849. In contrast, the AUC for the combined model with the peritumoral 4 mm region decreased from 0.904 to 0.782, while the AUC for the combined model including intratumoral features improved from 0.842 to 0.858. This may be attributed to the blurred boundary between NML and normal tissue, which hinders the clinical and ultrasound independent predictive factors from capturing the features of the peritumoral region for the combined model. Additionally, the variations in the peritumoral 4 mm region's imaging features are significantly different from those of the intratumoral features, and both the clinical and ultrasound models primarily represent intratumoral characteristics, resulting in poor complementarity and leading to inferior predictive performance for the combined model with the peritumoral 4 mm region compared to the original imaging model. In contrast, the peritumoral 2 mm region is more closely aligned with the peritumoral microenvironment, enhancing its complementarity with the clinical and ultrasound models. The constructed DCA curves indicate that the best predictive performance is achieved with the combined model that includes the peritumoral 2 mm region (AUDC: 0.294, 95%: -0.530 to -0.549), followed by the combined model with intratumoral features (AUDC: 0.282, 95%: -0.306 to -0.297). The DCA curve reveals that when the threshold probability is very low, the model's net benefit is negative. This is because the lower end of the DCA curve often overlaps with the net benefit curve of the "Treat-all" strategy, indicating that at very low thresholds, the model cannot provide greater net clinical benefit than the empirical treat-all strategy. Furthermore, the model's false positive predictions may lead to a large number of benign cases being misdiagnosed as malignant and undergoing unnecessary biopsies or surgeries, where the clinical harm outweighs the benefit. The results of this study demonstrate that the nomograms combining intratumoral and peritumoral 2 mm imaging scores with clinical and ultrasound independent predictive factors exhibit similar predictive performance in distinguishing benign and malignant NWL, with C-index values of 0.982 and 0.978, respectively. This suggests a high diagnostic value that can provide valuable reference for clinical decision-making.

This study has certain limitations: it is a single-center retrospective study, and non-mass breast cancer is relatively rare in clinical practice, resulting in a small sample size. The study utilized internal validation with cross-validation to assess the model's performance, but the lack of an external independent validation group may introduce a series of biases. The reliability and stability of these results need to be confirmed by increasing the sample size and conducting multicenter studies. Due to the fact that NWL does not meet the BI-RADS criteria for mass characteristics on ultrasound images, there are challenges in the annotation and contouring of two-dimensional ultrasound images. To address this issue, we chose to delineate the most prominent parts of the lesions. However, this approach makes the image segmentation susceptible to subjective differences among different observers. Future studies could consider employing automated segmentation methods to reduce inter-observer variability.

Data availability
Data is provided within the manuscript or supplementary information files.

Code availability
The related code is open-source and available in the GitHub repository: https://github.com/Jarrily/NML/tree/master and Zenodo https://doi.org/10.5281/zenodo.10421362 with https://doi.org/10.5281/zenodo.10421362.

References
1. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **74**, 229–263 (2024).
2. Zhang, J. et al. Re-evaluation of high-risk breast mammography lesions by target ultrasound and ABUS of breast non-mass-like lesions. *BMC Med Imaging* **21**, 156 (2021).
3. Magny, S. J., Shikhman, R., Keppke, A. L., in *StatPearls*. (StatPearls Publishing Copyright © 2024, StatPearls Publishing LLC., Treasure Island (FL), 2024).
4. Zhang, W. et al. Non-mass breast lesions on ultrasound: feature exploration and multimode ultrasonic diagnosis. *Ultrasound Med Biol* **44**, 1703–1711 (2018).
5. Yabuuchi, H. et al. Non-mass-like enhancement on contrast-enhanced breast MR imaging: lesion characterization using combination of dynamic contrast-enhanced and diffusion-weighted MR images. *Eur J Radiol* **75**, e126-132 (2010).
6. Brem, R. F., Lenihan, M. J., Lieberman, J. & Torrente, J. Screening breast ultrasound: past, present, and future. *AJR Am J Roentgenol* **204**, 234–240 (2015).
7. Wang, Z. L., Li, N., Li, M. & Wan, W. B. Non-mass-like lesions on breast ultrasound: Classification and correlation with histology. *Radiol Med* **120**, 905–910 (2015).
8. Bahl, M. Combining AI and radiomics to improve the accuracy of breast US. *Radiology* **312**, e241795 (2024).
9. Magnuska, Z. A. et al. Combining radiomics and autoencoders to distinguish benign and malignant breast tumors on US images. *Radiology* **312**, e232554 (2024).

10. Meng, Y., Yang, Y., Hu, M., Zhang, Z. & Zhou, X. Artificial intelligence-based radiomics in bone tumors: Technical advances and clinical application. *Semin Cancer Biol* **95**, 75–87 (2023).

11. Xu, Z., Wang, Y., Chen, M. & Zhang, Q. Multi-region radiomics for artificially intelligent diagnosis of breast cancer using multimodal ultrasound. *Comput Biol Med* **149**, 105920 (2022).

12. Paget, S. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev* **8**, 98–101 (1989).

13. Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* **366**, 883–892 (2012).

14. Tan, Y. et al. Additive value of texture analysis based on breast MRI for distinguishing between benign and malignant non-mass enhancement in premenopausal women. *BMC Med Imaging* **21**, 48 (2021).

15. Leong, L. T. et al. Dual-energy three-compartment breast imaging for compositional biomarkers to improve detection of malignant lesions. *Commun. Med.* **1**, 29 (2021).

16. Malkov, S. et al. Mammographic texture and risk of breast cancer by tumor type and estrogen receptor status. *Breast Cancer Res.* **18**, 122 (2016).

17. Williams, J. *et al.* Tumor cell-adipocyte gap junctions activate lipolysis and contribute to breast tumorigenesis. *bioRxiv*, 277939 (2024).

18. Hangiandreou, N. J., Stekel, S. F. & Tradup, D. J. Comprehensive clinical implementation of DICOM structured reporting across a radiology ultrasound practice: Lessons learned. *J Am Coll Radiol* **14**, 298–300 (2017).

19. Monteiro, E., Costa, C. & Oliveira, J. L. A De-identification pipeline for ultrasound medical images in DICOM format. *J Med Syst* **41**, 89 (2017).

20. Yushkevich, P. A. et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128 (2006).

21. van Griethuysen, J. J. M. et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* **77**, e104–e107 (2017).

22. Okuno, T. *et al.*, Usefulness of color Doppler and strain elastography adjunctive to B-mode ultrasonography in the diagnosis of non-mass abnormalities of the breast: results of the BC-07 multicenter study of 385 cases. *J. Med. Ultrason. (2001)*, (2024).

23. Kubota, K., Mori, M., Fujioka, T., Watanabe, K. & Ito, Y. Magnetic resonance imaging diagnosis of non-mass enhancement of the breast. *J Med Ultrason* **2001**(50), 361–366 (2023).

24. Ko, K. H., Jung, H. K., Kim, S. J., Kim, H. & Yoon, J. H. Potential role of shear-wave ultrasound elastography for the differential diagnosis of breast non-mass lesions: Preliminary report. *Eur Radiol* **24**, 305–311 (2014).

25. Choe, J., Chikarmane, S. A. & Giess, C. S. Nonmass findings at breast US: Definition, classifications, and differential diagnosis. *Radiographics* **40**, 326–335 (2020).

26. Aydin, H. The MRI characteristics of non-mass enhancement lesions of the breast: Associations with malignancy. *Br J Radiol* **92**, 20180464 (2019).

27. Desai, P. & Aggarwal, A. Breast cancer in women over 65 years: A review of screening and treatment options. *Clin Geriatr Med* **37**, 611–623 (2021).

28. Zhang, P., Kong, P., Liu, D., He, S. & Zhang, X. Non-mass-type ductal carcinoma in situ of the breast on ultrasound: Features and pathological analysis. *J Cancer Res Ther* **20**, 665–668 (2024).

29. Amitai, Y. et al. Can breast MRI accurately exclude malignancy in mammographic architectural distortion?. *Eur Radiol* **30**, 2751–2760 (2020).

30. Choudhery, S., Polley, E. & Conners, A. L. Assessment of MRI-detected lesions on screening tomosynthesis in patients with newly diagnosed breast cancer. *Clin Imaging* **59**, 50–55 (2020).

31. Ohuchi, N. et al. 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* **387**, 341–348 (2016).

32. Sun, Q. *et al.*, Deep learning vs. radiomics for predicting axillary lymph node metastasis of breast cancer using ultrasound images: Don't forget the peritumoral region. *Front. Oncol.* **10**, 53 (2020).

33. Guo, S. et al. Multiregional radiomic model for breast cancer diagnosis: Value of ultrasound-based peritumoral and parenchymal radiomics. *Quant Imaging Med Surg* **13**, 3127–3139 (2023).

34. Huang, Z. et al. Optimizing breast cancer diagnosis with photoacoustic imaging: An analysis of intratumoral and peritumoral radiomics. *Photoacoustics* **38**, 100606 (2024).

35. Liu, W., Li, L., Deng, J. & Li, W. A comprehensive approach for evaluating lymphovascular invasion in invasive breast cancer: Leveraging multimodal MRI findings, radiomics, and deep learning analysis of intra- and peritumoral regions. *Comput Med Imaging Graph* **116**, 102415 (2024).

36. Yu, T. et al. Integrating intratumoral and peritumoral radiomics with deep transfer learning for DCE-MRI breast lesion differentiation: A multicenter study comparing performance with radiologists. *Eur J Radiol* **177**, 111556 (2024).

37. Wang, J., Gao, W., Lu, M., Yao, X. & Yang, D. Development of an interpretable machine learning model for Ki-67 prediction in breast cancer using intratumoral and peritumoral ultrasound radiomics features. *Front Oncol* **13**, 1290313 (2023).

38. Bernatz, S. et al. CT-radiomics and clinical risk scores for response and overall survival prognostication in TACE HCC patients. *Sci Rep* **13**, 533 (2023).

39. Wang, Y. et al. Contrast-enhanced ultrasound-magnetic resonance imaging radiomics based model for predicting the biochemical recurrence of prostate cancer: A feasibility study. *Comput Math Methods Med* **2022**, 8090529 (2022).

Author contributions
J.L. and X.L. conceived and designed the full research. T.L., Y.L. and R.L. all collected case data. J.L. conducted traditional radiomics feature extraction, screening, and extreme model construction. J.L. and J.C. conducted analysis, interpretation of the model and chart plotting. J.L. drafted the manuscript. All authors contributed to revisions of the manuscript. The final version of the manuscript was read and approved by all authors. All authors had full access to the full data in the study and accept responsibility to submit for publication.

Funding

Declarations

Competing interests
The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-03704-2.

Correspondence and requests for materials should be addressed to X.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.