


# Multimodal Machine Learning-Based Ductal Carcinoma in situ Prediction from Breast Fibromatosis

Yan Jiang<sup>1</sup>, Yuanyuan Peng<sup>1</sup>, Yingyi Wu<sup>1</sup>, Qing Sun<sup>1</sup>, Tebo Hua<sup>1,2</sup> 

<sup>1</sup>Department of Ultrasound, Ningbo Medical Centre Lihuili Hospital, Ningbo, Zhejiang, People's Republic of China; <sup>2</sup>Department of Thyroid Breast Surgery, Ningbo Medical Centre Lihuili Hospital, Ningbo, Zhejiang, People's Republic of China

Correspondence: Tebo Hua, Department of Thyroid Breast Surgery, Ningbo Medical Centre Lihuili Hospital, Ningbo, Zhejiang, 315040, People's Republic of China, Email huatebo@163.com

**Objective:** To develop a clinical-radiomics model using a multimodal machine learning method for distinguishing ductal carcinoma in situ (DCIS) from breast fibromatosis.

**Methods:** The clinical factors, ultrasound features, and related ultrasound images of 306 patients (198 DCIS patients) were retrospectively collected. Patients in the development and validation cohort were 184 and 122, respectively. The independent clinical and ultrasound factors identified by the multivariable logistic regression analysis were used for the clinical-ultrasound model construction. Then, the region of interest of breast lesions was delineated and radiomics features were extracted. Six machine learning algorithms were trained to develop a radiomics model. The algorithm with higher and more stable prediction ability was chosen to convert the output of the results into the Radscore. Further, the independent clinical predictors and Radscore were enrolled into the logistic regression analysis to generate a combined clinical-radiomics model. The receiver operating characteristic curve analysis, DeLong test, and decision curve analysis were adopted to compare the prediction ability and clinical efficacy of three different models.

**Results:** Among the six classifiers, logistic regression model was selected as the final radiomics model. Besides, the combined clinical-radiomics model exhibited a superior ability in distinguishing DCIS from breast fibromatosis to the clinical-ultrasound model and the radiomics model.

**Conclusion:** The combined model by integrating clinical-ultrasound factors and radiomics features performed well in predicting DCIS, which might promote prompt interventions to improve the early diagnosis and prognosis of the patients.

**Keywords:** multimodal machine learning, clinical-ultrasound features, radiomics, ductal carcinoma in situ

## Introduction

Breast cancer is the most frequent malignant tumor in females worldwide with over half of these cases occurring in developed countries.<sup>1</sup> In the United States and Europe, breast cancer is one of the leading causes of cancer-related mortality.<sup>2</sup> Besides, breast cancer is diagnosed in 12.2% of women in China.<sup>3</sup> With medical advances, the calcifications and breast imaging reporting and data system (BI-RADS) lexicon have been identified which represent variable underlying histologic conditions related to ductal carcinoma in situ (DCIS) or invasive breast cancer.<sup>4</sup> It was estimated that about 51,400 females would have DCIS in 2022 according to the American Cancer Society.<sup>5</sup> At stage 0 of breast cancer, DCIS is a precursor to invasive ductal carcinoma and features a neoplastic proliferation of epithelial cells surrounded by myoepithelial cells. Myoepithelial cells are confined by an intact basement membrane that separates these cells from the breast stroma and prevents tumor cells from metastasizing.<sup>6</sup> Early detection can prevent DCIS from progressing into a more invasive cancer.<sup>6</sup> Breast fibromatosis is a locally invasive lesion with no metastatic potential caused by fibroblasts or myofibroblasts that may occur within the breast parenchyma or originate from the thoracic fascia and extend into the breast.<sup>7</sup> Therefore, it is necessary to investigate a novel method to identify DCIS or breast fibromatosis and give corresponding treatment to improve the diagnosis and prognosis of patients.

Ultrasound features including infiltrative margins, microcalcifications, and hypoechoic have been reported to be related to higher suspicion of malignancy.<sup>8</sup> To achieve a more quantitative prediction method, radiomics has been

introduced into many studies. Radiomics is involved in converting digital images into high-dimensional mineable features using high-throughput extraction analysis to explore the underlying tissue information of the images that are hard to recognize by the human eye.<sup>9</sup> Radiomics performs more satisfactorily in predicting multi-parameter imaging features than traditional imaging indicators. These extracted features can be used to build efficient models by incorporating with genomic, histologic, laboratory, and clinical data for prediction, histological classification or classification of benign and malignant lesions, lymph node metastasis detection, and diagnosis by using machine learning (ML) methods.<sup>10–13</sup> In addition, the predictive value of radiomics for treatment selection or treatment response and clinical prognosis has been demonstrated, which is closely associated with proteomic, transcriptomic, and genomic characteristics.<sup>14,15</sup> However, few studies developed ML models to distinguish DCIS from breast fibromatosis.

ML is an emerging artificial intelligence tool, critical in improving prediction accuracy of the diagnosis and prognosis.<sup>16</sup> Improved medical database management yielded a potential use of ML in medicine.<sup>17</sup> Compared to other statistical methods, ML algorithms allow interactions between variables, identify key predictors, find optimal algorithms between the study outcomes and potential predictors by learning from dataset patterns, and show greater accuracy in clinical settings.<sup>18</sup> Our study incorporated the clinical data, ultrasonic characteristics, and radiomics features to build related models. The predictive ability and clinical efficacy of the models were compared to select an optimal model for predicting DCIS.

## Materials and Methods

### Patients and Data Collection

Patients diagnosed with DCIS (n=198) via surgery and those diagnosed with breast fibromatosis (n=108) via B-mode ultrasound from Ningbo Medical Centre Lihuili Hospital between January 2016 and June 2023 were enrolled in the study. The 306 participants were randomly divided into the development cohort and internal validation cohort at a ratio of 6:4 (184 in the development cohort and 122 in the internal validation cohort).

Inclusion criteria: (1) Patients were confirmed as DCIS by postoperative pathology, and immunohistochemistry test results were obtained or those diagnosed with breast fibromatosis by B-mode ultrasound; (2) Breast ultrasound was performed within 1 week prior to surgery and the image quality was good; (3) There were no other primary tumors.

Exclusion criteria: (1) Poor image quality; (2) Patients who have previously undergone breast puncture, surgery, radiation, chemotherapy, or hormone therapy.

The clinical and B-ultrasonic features of patients were collected: age, tumor size, tumor location, morphology, calcification, blood flow, BI-RADS grade, orientation, margin, echo mode, and rear echo signature. The multivariate logistic regression analysis was utilized to select the independent clinical and B-ultrasonic features as a clinical-ultrasonic model to predict DCIS. The value of the clinical-ultrasonic model in predicting DCIS was analyzed by receiver operating characteristic (ROC) curve analysis in the training cohort, test cohort, and validation cohort.

Under the Declaration of Helsinki, this study involving human participants was reviewed and approved by the Ethics Committee of Ningbo Medical Centre Lihuili Hospital (Approval NO.KY2022PJ139). Written informed consent for participation was not required for this study following the national legislation and the institutional requirements.

### Image Acquisition and Segmentation

The ultrasound doctor performed a routine two-dimensional ultrasound examination on both mammary glands of the patient, identified the breast lesions, recorded the location, shape boundary, and blood flow, marked the lesions on the body surface, and scanned the bilateral axillary lymph nodes. The diagnostic report was obtained by referring to the BI-RADS classification standard (5th edition): PHILIPS IU22, Q5, MINDRAY DC-8, APLIO5000, probe frequency 3–12 MHz.

The sonographer selected the largest plane of each breast lesion without knowing the histopathological information and delineated a region of interest (ROI) covering the entire lesion with the ITK-SNAP software (open-source software; <http://www.itk-snap.org>).



## Radiomics Feature Extraction and Evaluation

The ultrasound radiomics features including first-order, gray-level co-occurrence matrix (glcm), gray-level dependence matrix (gldm), gray-level run length matrix (glrlm), gray-level size zone matrix (glszm), and neighbourhood gray-tone difference matrix (ngtgm) features were extracted using the “pyradimoics” package in the 3D slicer software (version 5.4.0) by the intraobserver and interobserver. The interclass correlation coefficient (ICC) was employed to evaluate the consistency of the extracted radiomics features. Features with ICC >0.70 demonstrated a good consistency of these characteristics and were used for subsequent feature selection.

## Radiomics Feature Selection

All the radiomics features were performed with z-score normalization before screening features. Then, the least absolute shrinkage and selection operator (LASSO) method was adopted to select the important radiomics features. Next, collinearity analysis was conducted and features with a variance inflation factor (VIF) value greater than 6 indicating high collinearity were removed. Subsequently, the top five features were identified using XGBoost and random forest (RF) methods, respectively. The intersection of the two methods was considered as the vital features and used for the following radiomics model construction.

## Radiomics Model Development and Performance

In the development cohort, 147 (80%) patients were randomly selected to train models and tune hyperparameters, and 37 (20%) patients to test the models. We used six machine learning algorithms to develop the models: XGBoost, logistic regression, RF, support vector machines (SVM), K Nearest Neighbors (KNN), and multilayer perceptron (MLP). Resampling was utilized to address class imbalance. The area under the ROC curve (AUC), the area under the precision-recall curve (AP), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1 scores, and Kappa value served as evaluation metrics to compare the performance of six algorithms in the train and test sets. AUC was used to obtain the optimal model and to assess the consistency of the models in the training and test sets. The classifier with the best performance in the two sets was regarded as the radiomics model and the predicted probability was considered as Radscore (RD). The ROC was used to explore the predictive performance of the model in distinguishing DCIS from fibroma patients in the training cohort, test cohort, and validation cohort.

## Clinical-RD Model Development

Factors in the clinical-ultrasonic model as well as the RD were enrolled into multivariate logistic regression analysis to select the independent predictors for clinical-RD construction. Then, the ROC was used to compare the predictive value of the clinical-ultrasonic model, the RD model, and the clinical-RD model for DCIS. Besides, decision curve analysis (DCA) was employed to evaluate the clinical usefulness of the three models in the development cohort and validation cohort, respectively.

## Statistical Analysis

All statistical analyses were performed using SPSS software (version 23.0) and R software (version 4.4.2). Categorical variables were expressed as count (percent) and the two-group differences were compared by Chi-square test or Fisher's exact test. After performing the Shapiro–Wilk test, all the continuous variables did not meet the normal distribution and were presented as median (quartiles). The group differences were analyzed by the Mann–Whitney–U test.  $P < 0.05$  indicated a significance level.

## Results

### Patient Characteristics and Clinical-Ultrasonic Model Construction

Patient characteristics in the development and validation cohorts are summarized in [Table 1](#). All the variables were not significant between the two cohorts. [Table 2](#) demonstrates the clinical characteristics and B-ultrasonic features in the development cohort. Patients had a median age of 37 and 52 years in the fibroma and DCIS groups, respectively ( $P$

**Table 1** Baseline Characteristics of Study Cohorts

Variables	Total (n=306)	Development Cohort (n=184)	Validation Cohort (n=122)	P-value
Age	46.71±13.71	46.28±13.27	47.37±14.32	0.497
Size	17.00 [11.00, 26.00]	18.00 [12.00, 28.00]	16.00 [10.00, 23.00]	0.223
Location				0.324
Inner upper	62 (20.25)	43 (23.37)	19 (15.58)	0.704
Inner lower	7 (2.29)	5 (2.72)	2 (1.64)	
Outer upper	192 (62.75)	109 (59.24)	83 (68.03)	
Outer lower	45 (14.71)	27 (14.67)	18 (14.75)	
Morphology				0.204
Ellipse	147 (48.04)	91 (49.46)	56 (45.90)	
Quasi-circle	35 (11.44)	19 (10.32)	16 (13.12)	
Irregularity	124 (40.52)	74 (40.22)	50 (40.98)	0.307
Calcification				
No	177 (57.84)	109 (59.24)	68 (55.74)	0.114
Thick and diffuse	20 (6.54)	15 (8.15)	5 (4.10)	
Dotted	109 (35.62)	60 (32.61)	49 (40.16)	
Blood flow				0.434
No	127 (41.50)	79 (42.94)	48 (39.34)	
Grade 1	117 (38.24)	67 (36.40)	50 (40.98)	
Grade 2	41 (13.40)	22 (11.96)	19 (15.58)	
Grade 3	21 (6.86)	16 (8.70)	5 (4.10)	0.376
BI-RADS grade				
<4A	96 (31.37)	64 (34.78)	32 (26.23)	
≥4A	210 (68.63)	120 (65.22)	90 (73.77)	0.970
Orientation				
Parallel	245 (80.07)	150 (81.52)	95 (77.87)	0.405
Unparallel	61 (19.93)	34 (18.48)	27 (22.13)	
Edge				
Regular	69 (22.55)	41 (22.28)	28 (22.95)	0.970
Lobular	109 (35.62)	71 (38.59)	38 (31.15)	
Irregularity	128 (41.83)	72 (39.13)	56 (45.90)	
Echo pattern				0.405
Uneven	58 (18.95)	35 (19.02)	23 (18.85)	
Cyst	248 (81.05)	149 (80.98)	99 (81.15)	
Rear echo signature				
Unaltered	243 (79.41)	149 (80.98)	94 (77.05)	
Altered	63 (20.59)	35 (19.02)	28 (22.95)	

**Abbreviation:** BI-RADS, breast imaging reporting and data system.

<0.05). In addition, tumor size is significantly larger in the DCIS group ( $P < 0.05$ ). The lesions in the DCIS group tended to exhibit irregularity, dotted calcification, higher blood flow grade, and BI-RADS grade ( $P < 0.05$ ). The edges were more likely to be angular and fuzzy in the DCIS group ( $P < 0.05$ ). The uneven echo pattern took a proportion of 10.45% and 23.93% in the fibroma and DCIS groups, respectively ( $P < 0.05$ ). The other factors including tumor location, orientation, and rear echo signature had no remarkably different distributions between the two groups ( $P > 0.05$ ).

Following this, the identified significant factors were included in multivariate logistic regression analysis. Morphology-quasi-circle, morphology-irregularity, calcification-thick, and diffuse, edge-leaflet, edge-angular and fuzzy, edge-cyst, and tumor size were not significantly related to DCIS ( $P > 0.05$ ), while calcification-dotted, BI-RADS grade, and age were independent factors for predicting DCIS ( $P < 0.05$ ) (Table 3). Thus, the three significant predictors were used for constructing a clinical-ultrasonic model. The AUCs for the clinical-ultrasonic model in predicting DCIS were 0.943, 0.938, and 0.907 respectively in the training, test, and validation cohorts (Figure 1).

**Table 2** Clinical Characteristics and B-Ultrasonic Features Between Fibroma and Carcinoma in situ Groups

Variables	Total (n=184)	Breast Fibromatosis (n=67)	Carcinoma in situ (n=117)	P-value
Age	46.00 [38.00, 55.00]	37.00 [27.00, 42.00]	52.00 [45.00, 59.00]	<0.001
Size	18.00 [12.00, 28.00]	15.00 [11.00, 21.00]	20.00 [12.00, 31.00]	0.010
Location				0.285
Inner upper	43 (23.37)	21 (31.34)	22 (18.80)	
Inner lower	5 (2.72)	2 (2.99)	3 (2.56)	
Outer upper	109 (59.24)	35 (52.24)	74 (63.25)	
Outer lower	27 (14.67)	9 (13.43)	18 (15.39)	
Morphology				<0.001
Ellipse	91 (49.46)	50 (74.63)	41 (35.04)	
Quasi-circle	19 (10.33)	8 (11.94)	11 (9.40)	
Irregularity	74 (40.21)	9 (13.43)	65 (55.56)	
Calcification				<0.001
No	109 (59.24)	61 (91.05)	48 (41.03)	
Thick and diffuse	15 (8.15)	4 (5.97)	11 (9.40)	
Dotted	60 (32.61)	2 (2.98)	58 (49.57)	
Blood flow				0.007
No	79 (42.94)	39 (58.21)	40 (34.19)	
Grade 1	67 (36.40)	21 (31.34)	46 (39.31)	
Grade 2	22 (11.96)	5 (7.46)	17 (14.53)	
Grade 3	16 (8.70)	2 (2.99)	14 (11.97)	
BI-RADS grade				<0.001
<4A	64 (34.78)	55 (82.09)	9 (7.69)	
≥4A	120 (65.22)	12 (17.91)	108 (92.31)	
Orientation				0.084
Parallel	150 (81.52)	59 (88.06)	91 (77.78)	
Unparallel	34 (18.48)	8 (11.94)	26 (22.22)	
Edge				<0.001
Regular	41 (22.28)	23 (34.33)	18 (15.39)	
Lobular	71 (38.59)	40 (59.70)	31 (26.50)	
Irregularity	72 (39.13)	4 (5.97)	68 (58.11)	
Echo pattern				0.025
Uneven	35 (19.02)	7 (10.45)	28 (23.93)	
Cyst	149 (80.98)	60 (89.55)	89 (76.07)	
Rear echo signature				0.624
Unaltered	149 (80.98)	53 (79.10)	96 (82.05)	
Altered	35 (19.02)	14 (20.90)	21 (17.95)	

**Abbreviation:** BI-RADS, breast imaging reporting and data system.

**Table 3** Association of Clinical-Ultrasonic Features with Carcinoma in situ

Variables	B	Odds Ratio (95% Confidence Interval)	P-value
Morphology-quasi-circle	-1.073	0.342 (0.056–2.908)	0.246
Morphology-irregularity	0.168	1.183 (0.238–5.889)	0.837
Calcification-thick and diffuse	-0.178	0.837 (0.133–5.257)	0.849
Calcification -dotted	2.671	14.448 (1.724–121.049)	0.014
Blood flow-grade 1	0.062	1.064 (0.238–4.752)	0.936

(Continued)

**Table 3** (Continued).

Variables	B	Odds Ratio (95% Confidence Interval)	P-value
Blood flow-grade 2	1.306	3.69 (0.282–48.236)	0.319
Blood flow-grade 3	−0.930	0.395 (0.020–7.593)	0.538
Edge-lobular	−0.725	0.484 (0.109–2.159)	0.342
Edge-irregularity	0.962	2.617 (0.331–20.661)	0.361
Echo pattern-cyst	−1.468	0.230 (0.041–1.309)	0.098
BI-RADS grade≥4A	2.813	16.654 (3.706–74.836)	<0.001
Age	0.109	1.115 (1.051–1.183)	<0.001
Size	0.032	1.033 (0.961–1.110)	0.383

**Abbreviation:** BI-RADS, breast imaging reporting and data system.

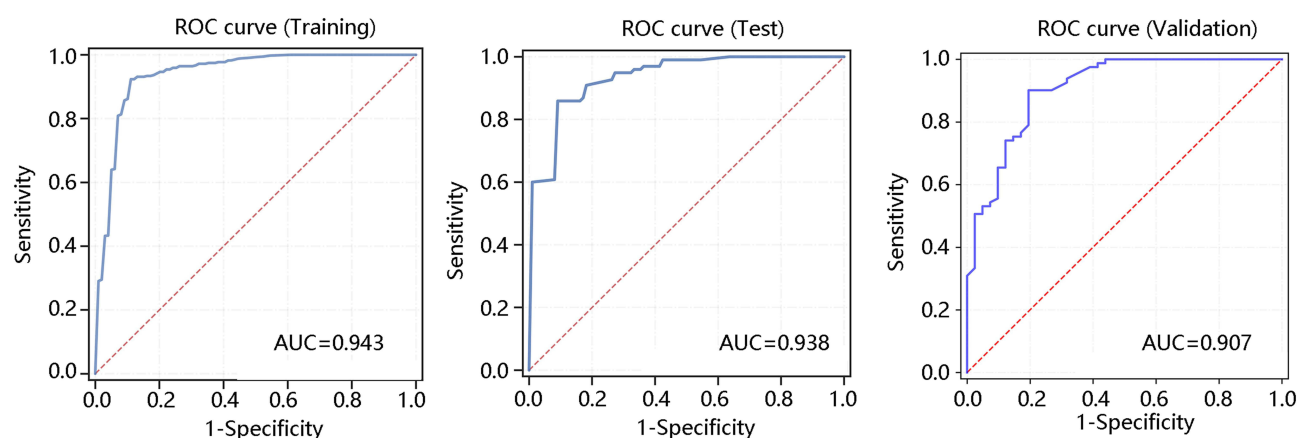
## Selection of Radiomics Features and Radiomics Model Construction

After LASSO screening, totally 14 features with non-zero coefficients were obtained (Figure 2A). The detailed information of these features is shown in Table 4. Further collinearity analysis showed that the VIF values of 10Percentile and interquartileRange were greater than 6, which were deleted for the following analysis (Table 5). Among the remaining 12 features, XGBoost and RF algorithm were respectively used to screen the top five important features (Figure 2B). Complexity and HighGrayLevelZoneEmphasis were the overlapped features of the two algorithms, which were used for constructing the radiomics model and its predicted probability called the RD.

The AUCs for XGBoost, logistic regression, RF, SVM, KNN, and MLP were 1.000, 0.976, 1.000, 0.989, 0.993, and 0.838, respectively in the training set and their AUCs were 0.961, 0.991, 0.986, 0.952, 0.974, and 0.867, respectively in the test set (Figure 3A). The APs for XGBoost, logistic regression, RF, SVM, KNN, and MLP were 1.000, 0.986, 1.000, 0.993, 0.994, and 0.860, respectively in the training set. The APs were 0.974, 0.997, 0.993, 0.971, 0.979, and 0.931, respectively in the test set (Figure 3B). The Results of other evaluation indexes for six machine learning methods are exhibited in Table 6. XGBoost was most likely to be susceptible to overfitting, whereas logistic regression was likely to be relatively stable. Thus, logistic regression was chosen as the final optimal model, which still owned satisfactory performance in the validation cohort (AUC=0.947) (Figure S1).

## Clinical-RD Model Generation and Evaluation

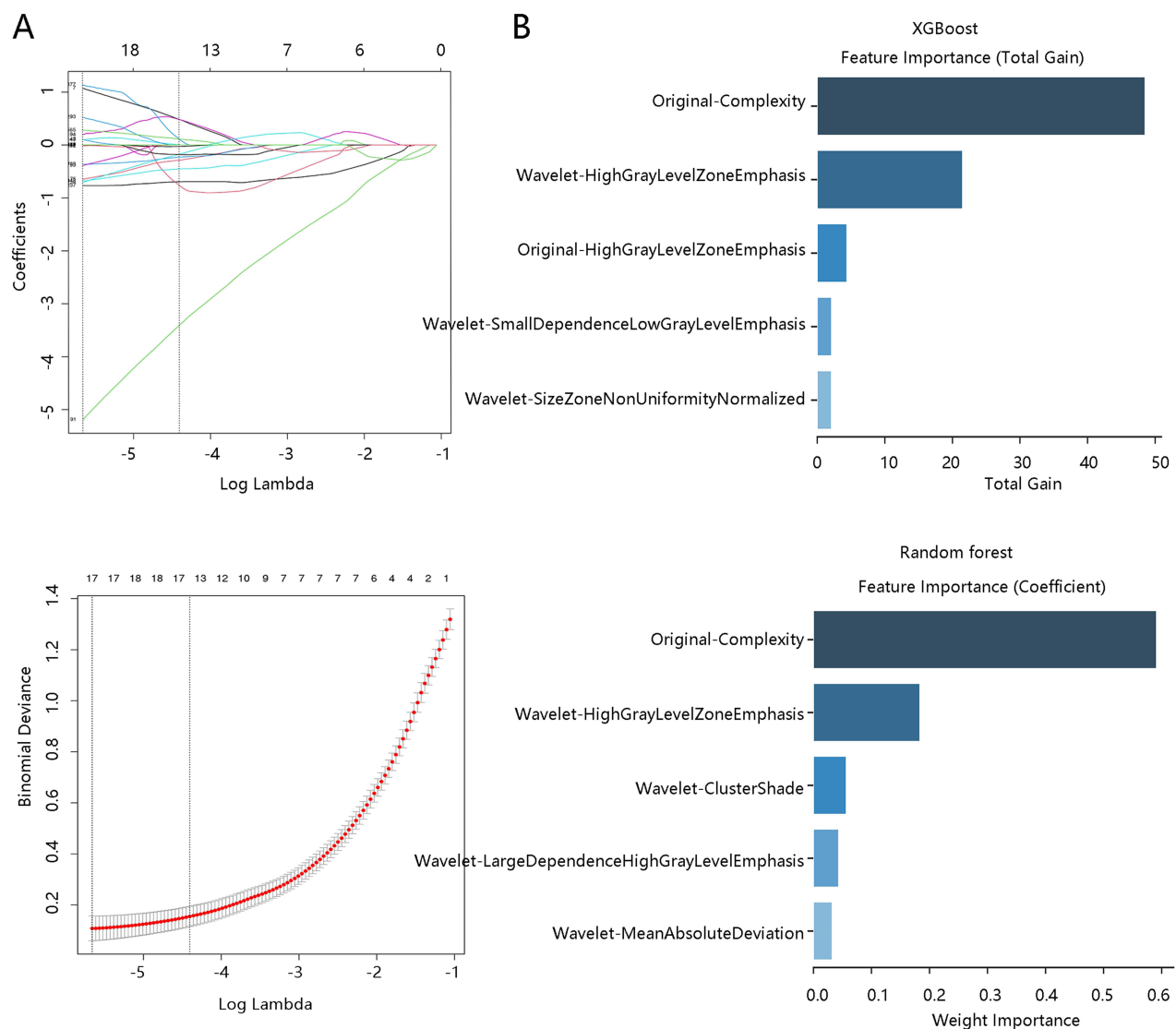
To obtain a more personalized prediction model, the variables in the clinical-ultrasonic and RD were included in multivariate logistic regression analysis. The results showed that RD, age, and BI-RADS grade were associated with



**Figure 1** The value of clinical-ultrasonic model in predicting carcinoma in situ.

**Abbreviations:** ROC, receiver operating characteristic; AUC, the area under the curve.





**Figure 2** Selection of radiomics features. (A) The radiomics features screened by LASSO. (B) The top five radiomics features were ranked by XGBoost and random forest algorithm.

DCIS with statistical significance and applied for clinical-RD generation ( $P < 0.05$ ) (Table 7). Finally, the clinical value of the three models was compared by the ROC and DCA in the development and validation cohorts. As shown in Figure 4A, clinical-RD model (AUC=0.996) had a superior ability to clinical-ultrasonic model (AUC=0.947) and radiomics model (AUC=0.980) in the development cohort with statistical difference ( $P < 0.05$ ) (Table 8). The same results were observed in the validation cohort with the AUCs of 0.983, 0.945, and 0.907 for the three models (Figure 4B and Table 8). The DCA results also showed that the clinical-RD model had the highest clinical efficacy compared to the radiomics model and the clinical-ultrasonic model in both development and validation cohorts (Figure 4C and D).

## Discussion

In this study, the clinical and ultrasonic features of patients with DCIS and breast fibromatosis were collected and analyzed. The factors that could independently predict the DCIS were enrolled in the clinical-ultrasonic model construction. Besides, the ultrasound images were collected to extract a large number of radiomics features. Based on the LASSO method, collinearity analysis, and machine learning algorithm, significant radiomics features were selected

**Table 4** Detailed Information of Selected Features with Non-Zero Coefficients

Image Type	Feature Class	Feature Name	Feature Coefficient
Original	firstorder	Maximum	0.486
Original	glszm	GrayLevelVariance	-0.827
Original	ngtdm	Complexity	-3.408
Wavelet-LLH	gldm	SmallDependenceLowGrayLevelEmphasis	0.004
Wavelet-LHL	firstorder	MeanAbsoluteDeviation	0.484
Wavelet-LHL	gldm	LargeDependenceHighGrayLevelEmphasis	-0.695
Wavelet-HLL	firstorder	10Percentile	-0.761
Wavelet-HLL	firstorder	InterquartileRange	0.113
Wavelet-HLL	glszm	HighGrayLevelZoneEmphasis	-0.453
Wavelet-HLH	glszm	SmallAreaLowGrayLevelEmphasis	-0.028
Wavelet-HHL	firstorder	Maximum	0.114
Wavelet-HHH	glszm	SizeZoneNonUniformityNormalized	-0.23
Wavelet-HHH	glszm	SmallAreaLowGrayLevelEmphasis	-0.176
Wavelet-LLL	glcm	ClusterShade	-0.179

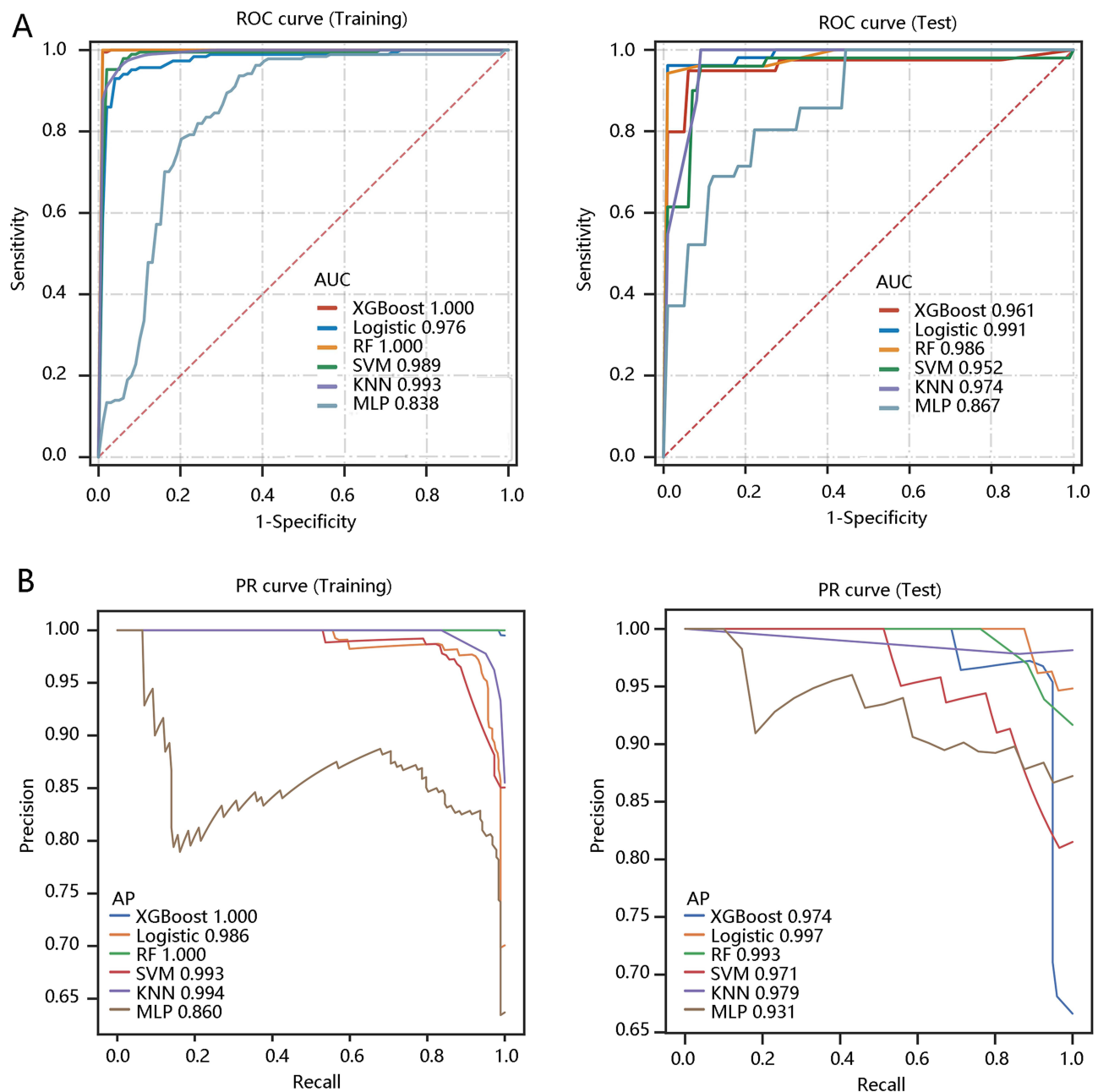
**Table 5** Collinearity Analysis

Variables	VIF
10Percentile	22.763
InterquartileRange	21.097
HighGrayLevelZoneEmphasis	5.709
LargeDependenceHighGrayLevelEmphasis	3.909
Complexity	3.802
HighGrayLevelZoneEmphasis	3.263
MeanAbsoluteDeviation	2.865
ClusterShade	2.607
Maximum	2.296
SizeZoneNonUniformityNormalized	2.288
SmallAreaLowGrayLevelEmphasis	2.188
Maximum	2.17
SmallAreaLowGrayLevelEmphasis	1.761
SmallDependenceLowGrayLevelEmphasis	1.441

for the radiomics model generation. Finally, the predictive value and clinical efficacy of the clinical-ultrasonic model, radiomics model, and the combined model were compared, among which the combined model exhibited superior ability.

Routine ultrasound is a basic imaging technique for detecting and diagnosing breast lesions.<sup>19</sup> Fine linear branching and fine pleomorphic calcifications are closely associated with malignancy of DCIS as well as an elevated risk of recurrence than other calcification morphologies.<sup>20</sup> In the BI-RADS atlas (second edition), breast lesions are ultimately assigned a grade after analyzing ultrasound characteristics. Grade 0: a diagnosis that requires a combination of other imaging tests. Grade 1: negative results or no lesions. Grade 2: benign lesion, no suspicious features. Grade 3: benign lesions with a 2% probability of malignancy. Grade 4: suspicious lesion with 2–95% probability of malignancy, biopsy is recommended. Due to the wide range of the probability, Grade 4 is further divided into 4A, 4B, and 4C with a malignant probability of 2–10%, 10–50%, and 50–95%, respectively. Grade 5: suspicious lesion with over 95% probability of malignancy. Grade 6: malignancy proved by biopsy.<sup>21</sup> Cha et al, reported that calcification was associated with pathological and biological markers.<sup>22</sup> Besides, Sharma et al, showed the importance of age in DCIS.<sup>23</sup> Herein, calcification-dotted, BI-RADS grade  $\geq 4A$  and older age were independently related to a higher probability of developing





**Figure 3** Predictive performance of six machine learning models. **(A)** Receiver operating characteristic curve analysis of the six classifiers in the training and test sets. **(B)** Precision curve analysis of the six classifiers in the training and test sets.

**Abbreviations:** ROC, receiver operating characteristic; AUC, the area under the curve; PR, precision curve; AP, area under the precision recall curve.

DCIS, and the three factors were used for building a clinical-ultrasonic model with satisfactory performance in distinguishing DCIS from breast fibromatosis.

As a method of quantitatively analyzing the gray value of medical images, radiomics reflects the texture and morphological characteristics of tumors.<sup>24</sup> Many quantitative features from medical images can be extracted by computer algorithms.<sup>25</sup> The degree of tumor heterogeneity is a prognostic biomarker related to cancer progression revealed by genomic analyses.<sup>25,26</sup> Notably, radiomics features characterize the heterogeneity throughout the entire tumor volume; however, biopsies only represent a small tumor portion or even a single site. Thus, radiomics features are promising markers of cancer aggressiveness.<sup>15</sup> In addition, radiomics features represent an essential role in predicting treatment response and patient survival.<sup>14</sup> After LASSO screening, and collinearity analysis, 12 radiomics features were identified

**Table 6** Predictive Performance of Six Machining Learning Classifiers in the Training and Test Sets

Model	Set	AUC (SD)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 Score	Kappa
XGBoost	Training	1.000 (0.000)	0.990	0.995	1.000	1.000	0.971	0.997	0.977
	Test	0.961 (0.024)	0.878	0.949	0.971	1.000	0.804	0.974	0.761
Logistic regression	Training	0.976 (0.004)	0.939	0.941	0.954	0.972	0.889	0.956	0.871
	Test	0.991 (0.009)	0.946	0.962	1.000	0.979	0.893	0.969	0.881
Random forest	Training	1.000 (0.000)	0.993	1.000	1.000	1.000	0.982	1.000	0.986
	Test	0.986 (0.014)	0.959	0.940	1.000	0.979	0.923	0.959	0.909
Support vector machine	Training	0.989 (0.000)	0.963	0.974	0.963	0.979	0.937	0.976	0.919
	Test	0.952 (0.012)	0.905	0.960	0.927	0.919	0.891	0.939	0.789
K-NearestNeighbor	Training	0.993 (0.001)	0.929	0.962	0.955	0.984	0.864	0.972	0.853
	Test	0.974 (0.026)	0.932	1.000	0.955	0.978	0.857	0.989	0.850
Multilayer Perceptron	Training	0.838 (0.016)	0.844	0.973	0.636	0.821	0.908	0.890	0.639
	Test	0.867 (0.101)	0.824	0.950	0.748	0.816	0.813	0.878	0.565

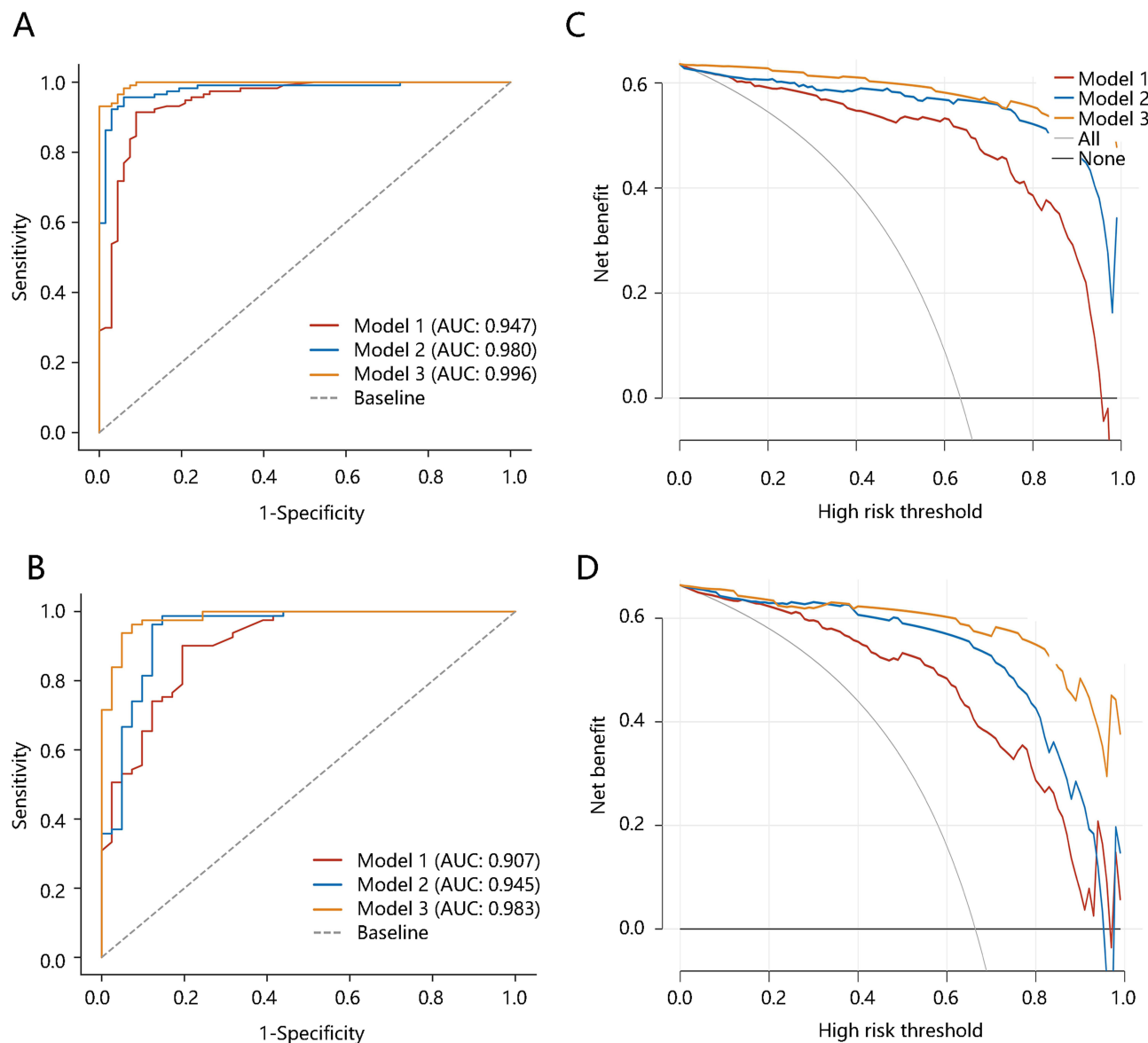
**Abbreviations:** AUC, area under the curve; PPV, Positive predictive value; NPV, negative predictive value.

**Table 7** Logistic Regression Analysis Results

Variables	B	Odds ratio (95% Confidence Interval)	P-value
Radscore	0.275	1.317 (1.175–1.603)	<0.001
Age	0.132	1.141 (1.036–1.307)	0.020
BI-RADS grade≥4A	3.626	37.575 (4.718–621.76)	0.002
Calcification-thick and diffuse	−0.277	0.758 (0.048–12.998)	0.839
Calcification-dotted	0.585	1.795 (0.118–45.385)	0.673

and their feature classes included firstorder, glszm, ngtdm, gldm, and glcm. The First-order features including gray-level mean, maximum, minimum, variance, and percentiles are based on the global gray-level histogram.<sup>9</sup> GLSZM counts the number of zones of linked voxels with the same gray level from the basis of the matrix.<sup>27</sup> NGTDM (Neighborhood Gray-Tone Difference Matrix) quantifies the difference between a gray level and the average gray level of its neighbors within a predefined distance.<sup>28</sup> GLDM (Gray Level Dependence Matrix) reflects the gray homogeneity of the local focus.<sup>29</sup> GLCM refers to the spatial relationship of pairs of pixels or voxels with predefined grayscale intensities and distances in different directions.<sup>27</sup> These traits that are impossible to detect by visual inspection can be exploited by radiomics, providing critical data for the prediction and diagnosis of diseases.<sup>30</sup> Interestingly, the logistic regression-based radiomics model outperformed the other models with good and stable prediction accuracy. Logistic regression is employed to assess the relationship between the binary dependent variables and independent variables, which can estimate the probability of the target variable. RF is an ensemble learning algorithm that uses multiple decision trees to provide a class prediction. SVM is a supervised model that identifies the optimal hyperplane, dividing the data into different classes. KNN is also a non-parametric learning algorithm in which output objects are classified using their local neighborhoods to formulate predictions. As a supervised learning algorithm, MLP can learn non-linear models. XG Boost uses an ensemble of weak prediction models for classification.<sup>31,32</sup>

Radiomics as a non-invasive method has been widely used in various diseases. Jin et al, reported that the ultrasound-based radiomics model incorporating clinical information and RD can effectively predict central lymph node metastasis in papillary thyroid carcinoma patients with Hashimoto's thyroiditis.<sup>33</sup> The machine learning model based on ultrasound radiomics features performed well in discriminating primary and metastatic liver cancer.<sup>34</sup> Besides, the clinical parameters combined with the radiomics can independently predict the microvascular invasion in hepatocellular carcinoma.<sup>35</sup> The MRI-based radiomics model can predict the overall survival and tumor-infiltrating macrophages in gliomas.<sup>36</sup> This



**Figure 4** Predictive value and clinical efficacy of the three models in the development and validation cohorts. Receiver operating characteristic curve analysis of three models in the (A) development cohort and (B) validation cohort. Decision curve analysis of the three models in the (C) development cohort and (D) validation cohort. Model 1: clinical-ultrasonic model; model 2: radiomics model; model 3: Clinical-radiomics model.

**Abbreviation:** AUC, area under the curve.

study revealed the significance of the model combined with the independent clinical-ultrasonic features and the RD obtained from the radiomics in distinguishing DCIS from breast fibromatosis.

For strengths, our study adopted the multimodal machine learning method to construct different models and selected the optimal classifier for distinguishing DCIS from breast fibromatosis. Then, the predictive value and clinical efficacy of the three models were explored in the development and validation cohorts using ROC and DCA. Despite some promising findings, only the ultrasound images were collected. The magnetic resonance imaging and mammography will be conducted to extract relevant radiomics features and the clinical efficacy of different radiomics models based on different examination methods should be compared in the future.

In conclusion, our study employed machine learning methods to develop a model effectively predicting the DCIS. The combined model achieved a satisfactory performance in distinguishing DCIS from breast fibromatosis.

**Table 8** The DeLong Test for Comparing the Area Under the Curves of Three Models

Models	Clinical-Ultrasonic Model	Radiomics Model	Clinical-RD Model
<b>Development cohort</b>			
Clinical-ultrasonic model	/	0.109	0.004
Radiomics model	0.109	/	0.048
Clinical-RD model	0.004	0.048	/
<b>Validation cohort</b>			
Clinical-ultrasonic model	/	0.327	0.002
Radiomics model	0.327	/	0.134
Clinical-RD model	0.002	0.134	/

## Ethics Approval and Consent to Participate

Under the Declaration of Helsinki, this study involving human participants were reviewed and approved by the Ethics Committee of Ningbo Medical Centre Lihuili Hospital (Approval NO.KY2022PJ139). Our research conforms to patient data confidentiality and compliance with the Declaration of Helsinki.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This work was supported by the Ningbo Science and Technology Programme for Public Welfare (2022S038) and Natural Science Foundation Program of Ningbo (2023J230). The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Disclosure

The authors have no conflicts of interest to disclose in this work.

## References

1. Liu Y, Zhao S, Zhang Y, Onwuka JU, Zhang Q, Liu X. Bisphosphonates and breast cancer survival: a meta-analysis and trial sequential analysis of 81508 participants from 23 prospective epidemiological studies. *Aging*. 2021;13(15):19835–19866. doi:10.18632/aging.203395
2. Manley H, Mutasa S, Chang P, Desperito E, Crew K, Ha R. Dynamic changes of convolutional neural network-based mammographic breast cancer risk score among women undergoing chemoprevention treatment. *Clin Breast Cancer*. 2021;21(4):e312–e318. doi:10.1016/j.clbc.2020.11.007
3. Li J, Song Y, Xu S, et al. Predicting underestimation of ductal carcinoma in situ: a comparison between radiomics and conventional approaches. *Int J Comput Assist Radiol Surg*. 2019;14(4):709–721. doi:10.1007/s11548-018-1900-x
4. Venkatesan A, Chu P, Kerlikowske K, Sickles EA, Smith-Bindman R. Positive predictive value of specific mammographic findings according to reader and patient variables. *Radiology*. 2009;250(3):648–657. doi:10.1148/radiol.2503080541
5. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin*. 2022;72(1):7–33. doi:10.3322/caac.21708
6. Hophan SL, Odnokoz O, Liu H, et al. Ductal carcinoma in situ of breast: from molecular etiology to therapeutic management. *Endocrinology*. 2022;163(4). doi:10.1210/endocr/bqac027
7. Liu H, Zeng H, Zhang H, et al. Breast fibromatosis: imaging and clinical findings. *Breast J*. 2020;26(11):2217–2222. doi:10.1111/tbj.14008
8. Friedrich-Rust M, Meyer G, Dauth N, et al. Interobserver agreement of thyroid imaging reporting and data system (TIRADS) and strain elastography for the assessment of thyroid nodules. *PLoS One*. 2013;8(10):e77927. doi:10.1371/journal.pone.0077927
9. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to Radiomics. *J Nucl Med*. 2020;61(4):488–495. doi:10.2967/jnumed.118.222893
10. Wu J, Fang Q, Yao J, et al. Integration of ultrasound radiomics features and clinical factors: a nomogram model for identifying the Ki-67 status in patients with breast carcinoma. *Front Oncol*. 2022;12:979358. doi:10.3389/fonc.2022.979358
11. Wang X, Agyekum EA, Ren Y, et al. A radiomic nomogram for the ultrasound-based evaluation of extrathyroidal extension in papillary thyroid carcinoma. *Front Oncol*. 2021;11:625646. doi:10.3389/fonc.2021.625646

12. Zhou SC, Liu TT, Zhou J, et al. An ultrasound radiomics nomogram for preoperative prediction of central neck lymph node metastasis in papillary thyroid carcinoma. *Front Oncol.* **2020**;10:1591. doi:10.3389/fonc.2020.01591
13. Wang R, Dai W, Gong J, et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. *J Hematol Oncol.* **2022**;15(1):11. doi:10.1186/s13045-022-01225-3
14. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* **2016**;278(2):563–577. doi:10.1148/radiol.2015151169
15. Sala E, Mema E, Himoto Y, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol.* **2017**;72(1):3–10. doi:10.1016/j.crad.2016.09.013
16. Zou Y, Shi Y, Liu J, et al. A comparative analysis of six machine learning models based on ultrasound to distinguish the possibility of central cervical lymph node metastasis in patients with papillary thyroid carcinoma. *Front Oncol.* **2021**;11:656127. doi:10.3389/fonc.2021.656127
17. Masuda T, Nakaura T, Funama Y, et al. Machine learning to identify lymph node metastasis from thyroid cancer in patients undergoing contrast-enhanced CT studies. *Radiography.* **2021**;27(3):920–926. doi:10.1016/j.radi.2021.03.001
18. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol.* **2017**;2(2):204–209. doi:10.1001/jamacardio.2016.3956
19. Rahbar G, Sie AC, Hansen GC, et al. Benign versus malignant solid breast masses: US differentiation. *Radiology.* **1999**;213(3):889–894. doi:10.1148/radiology.213.3.r99dc20889
20. Woodard GA, Price ER. Qualitative radiogenomics: association between BI-RADS calcification descriptors and recurrence risk as assessed by the oncotype DX ductal carcinoma in situ score. *AJR Am J Roentgenol.* **2019**;212(4):919–924. doi:10.2214/AJR.18.20306
21. 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(1):11921. doi:10.1038/s41598-019-48488-4
22. Cha H, Chang YW, Lee EJ, et al. Ultrasonographic features of pure ductal carcinoma in situ of the breast: correlations with pathologic features and biological markers. *Ultrasonography.* **2018**;37(4):307–314. doi:10.14366/uscg.17039
23. Radswiki SR, Ashraf A. Ductal carcinoma in situ. The Radswiki; **2010**.
24. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol.* **2017**;28(6):1191–1206. doi:10.1093/annonc/mdx034
25. Braman NM, Etesami M, Prasanna P, et al. 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(1):57. doi:10.1186/s13058-017-0846-1
26. Morris LG, Riaz N, Desrichard A, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget.* **2016**;7(9):10051–10063. doi:10.18632/oncotarget.7067
27. Yu Q, Liu J, Lin H, Lei P, Fan B. Application of radiomics model of CT images in the identification of ureteral calculus and phlebolith. *Int J Clin Pract.* **2022**;2022:5478908. doi:10.1155/2022/5478908
28. Gu S, Qian J, Yang L, et al. Multiparametric MRI radiomics for the differentiation of brain glial cell hyperplasia from low-grade glioma. *BMC Med Imaging.* **2023**;23(1):116. doi:10.1186/s12880-023-01086-3
29. Tian H, Wu H, Wu G, Xu G. Noninvasive prediction of TERT promoter mutations in high-grade glioma by radiomics analysis based on multiparameter MRI. *Biomed Res Int.* **2020**;2020:3872314. doi:10.1155/2020/3872314
30. Sasaki T, Kinoshita M, Fujita K, et al. Radiomics and MGMT promoter methylation for prognostication of newly diagnosed glioblastoma. *Sci Rep.* **2019**;9(1):14435. doi:10.1038/s41598-019-50849-y
31. Salman R, Alzaatreh A, Sulieman H, Faisal S. A bootstrap framework for aggregating within and between feature selection methods. *Entropy.* **2021**;23(2):200. doi:10.3390/e23020200
32. Wang X, Wang Y, Xu Z, Xiong Y, Wei DQ. ATC-NLSP: prediction of the classes of anatomical therapeutic chemicals using a network-based label space partition method. *Front Pharmacol.* **2019**;10:971. doi:10.3389/fphar.2019.00971
33. Jin P, Chen J, Dong Y, et al. Ultrasound-based radiomics nomogram combined with clinical features for the prediction of central lymph node metastasis in papillary thyroid carcinoma patients with Hashimoto's thyroiditis. *Front Endocrinol.* **2022**;13:993564. doi:10.3389/fendo.2022.993564
34. Mao B, Ma J, Duan S, Xia Y, Tao Y, Zhang L. Preoperative classification of primary and metastatic liver cancer via machine learning-based ultrasound radiomics. *Eur Radiol.* **2021**;31(7):4576–4586. doi:10.1007/s00330-020-07562-6
35. Hu HT, Wang Z, Huang XW, et al. Ultrasound-based radiomics score: a potential biomarker for the prediction of microvascular invasion in hepatocellular carcinoma. *Eur Radiol.* **2019**;29(6):2890–2901. doi:10.1007/s00330-018-5797-0
36. Li G, Li L, Li Y, et al. An MRI radiomics approach to predict survival and tumour-infiltrating macrophages in gliomas. *Brain.* **2022**;145(3):1151–1161. doi:10.1093/brain/awab340

## Cancer Management and Research

Dovepress

## Publish your work in this journal

Cancer Management and Research is an international, peer-reviewed open access journal focusing on cancer research and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/cancer-management-and-research-journal>