

Mammography Based Nomogram Integrating Radiomics and Clinical Features to Predict Benign or Malignant Regression of BI-RADS 4A Lesions at Follow-Up

Guoyan Yao^{1,*}, Lijun Chen^{1,*}, Tingfan Wu^{2,*}, Yuanyuan Liu¹, Yun Wan¹, Ziqiang Xia¹, Bo Liu¹, MinAn Zheng³

¹Department of Radiology, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, 510120, People's Republic of China; ²Central Research Institute, United Imaging Healthcare Group Co., Ltd, Shanghai, 210800, People's Republic of China; ³Department of Emergency, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, 510120, People's Republic of China

*These authors contributed equally to this work

Correspondence: Bo Liu, Department of Radiology, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, No. 111, Dade Road, Guangzhou, Guangdong Province, 510120, People's Republic of China, Email liubogzcm@163.com; MinAn Zheng, Department of Emergency, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, No. 111, Dade Road, Guangzhou, Guangdong Province, 510120, People's Republic of China, Email zhengminan@yeah.net

Purpose: The study aimed to develop a nomogram based on mammography radiomic and clinical features to predict the benign and malignant progression of BI-RADS 4a lesions under follow-up.

Materials and Methods: The retrospective study included 104 patients with over six months of follow-up, consisting of 56 malignant and 48 benign cases, totaling 202 images. Patients were randomly divided into training and validation sets at a 7:3 ratio. In total, 1316 radiomic features were extracted using AK3.30 software including morphological, first-order statistics and texture features. Spearman correlation analysis and the least absolute shrinkage and selection operator (LASSO) method were performed for feature selection. Univariate and multivariate logistic regression analyses were used to identify independent risk factors among clinical features and construct a radiomic-clinical fusion nomogram. The performance of radiomics model and radiomic-clinical fusion model were evaluated using the area under the receiver operating characteristic (ROC) curve. DeLong test was employed to compare the efficacy between the two models.

Results: Four radiomic features were selected, combined with two clinical features (positive clinical palpation and history of breast surgery). The AUC values for the radiomics model and radiomic-clinical fusion model in the training and testing groups were 0.858 and 0.860, and 0.923 and 0.904, respectively. The DeLong test showed no significant difference between the two models with a *P* value > 0.05.

Conclusion: The nomogram based on mammography radiomics and clinical features demonstrated good performance in predicting the benign and malignant progression of BI-RADS 4a lesions under follow-up, showing potential for risk stratification of BI-RADS 4a lesions.

Keywords: BI-RADS 4a, radiomics, nomogram, clinical characteristic, predictive model

Introduction

Breast cancer is the leading threat to women's health globally, and its disease burden continues to grow. According to GLOBOCAN 2022, breast cancer is the second most prevalent malignant tumour in women with 2.3 million new cases per year, and the fourth leading cause of cancer-related deaths with 670,000 deaths.¹ In early screening systems, mammography in combination with the BI-RADS (Breast Imaging Reporting and Data System) classification criteria

has become a core strategy for reducing breast cancer mortality.² Mammography demonstrated significant value in breast screening and lesion detection owing to the advantage of high-resolution imaging, which could clearly display the structures of the breast gland, ducts and other tissues (10.1186/s13244-025-01994-8; 10.1007/s00432-025-06122-z). According to the latest guidelines of the American College of Radiology (ACR), BI-RADS category 4A lesions are defined as suspicious lesions with a 2%–10% risk of malignancy, which need to be confirmed by tissue biopsy.^{3,4}

Available clinical practice suggests significant limitations in the diagnostic efficacy of the BI-RADS 4A classification. A National Mammography Database (NMD) analysis of 16,784 cases with a preoperative classification of 4A and biopsy showed that only 7.6% (1274) were ultimately diagnosed with malignant.⁵ This high false-positive rate would subject many patients with benign lesions to unnecessary invasive biopsy procedures which may cause hematoma or patient discomfort, delay prioritisation of high-risk patients and increase the workload of the pathology department.⁶ Therefore, the need to accurately identify true malignant cases in category 4A lesions has become an urgent need to minimize patient harm while optimizing the utilization of healthcare resources for breast cancer screening.

In recent years, with the advancement of radiomics technology, it has shown tremendous potential in the diagnosis and prediction of breast lesions. Radiomics method extracts a vast array of quantitative information including shape, first-order statistics, texture and filtered features from medical imaging and modeling based on significant feature using machine-learning algorithms, which enables revelation of microscopic and molecular-level information of lesions and thereby enhancing the accuracy of benign and malignant discrimination⁷ (10.1148/radiol.2015151169). It has demonstrated promising results in the differentiation of benign and malignant breast lesions, prediction of breast cancer molecular typing and the efficacy of neoadjuvant chemotherapy.^{8–10} Previous studies have used radiomics based on MRI and ultrasound to distinguish between benign and malignant BI-RADS category 4 lesions and have achieved good results (10.1038/s41598-024-83347-x; 10.1186/s13014-025-02605-y; 10.21037/qims-24-1893). Nevertheless, Radiomics research on BI-RADS 4A based on Mammography remains scarce, with a focus on the analysis of single-modality imaging data and a lack of comprehensive utilization of multimodal data.¹¹ Additionally, existing studies still face challenges in clinical application, such as insufficient model generalizability and limited predictive accuracy.¹² Zhang et al (10.1007/s00330-023-09513-3) established a nomogram via intratumoral and peritumoral radiomics features and clinical risk factors, which showed the best performance in distinguishing benign and malignant BiRADS 4 lesions. However, to the best of our knowledge, no prior studies have combined radiomics features and clinical predictors for benign or malignant regression of BI-RADS 4A lesions at follow-up.¹³

This study aims to construct a fusion Nomogram model based on radiomics features from mammographic images and clinical characteristics to predict the benign and malignant progression of BI-RADS 4A lesions. The expected results will provide clinicians with quantitative decision support tools to reduce non-essential biopsies under the premise of ensuring diagnostic sensitivity, and ultimately promote the application of precision medicine in breast cancer screening.

Materials and Methods

Study Design and Ethical Review

The study was a single-centre retrospective cohort study approved by the Ethics Committee of The Second Affiliated Hospital of Guangzhou University of Chinese Medicine (approval number: ZE2021-111), following the ethical principles of the Declaration of Helsinki on medical research. Informed consent requirements were waived due to the use of de-identified historical data.

Study Population

Inclusion criteria were as follows: (1) patients received ≥ 2 mammograms with a time interval ≥ 6 months between January 2010 and October 2020; (2) patients evaluated as presence of BI-RADS category 4A lesions on baseline examination; (3) final examination showed imaging progression ($\geq 20\%$ increase in lesion size or increased morphological signs of malignancy); (4) gold standard for surgical resection pathology and complete clinical information (including age, reproductive history, family history of tumour, history of breast surgery and palpation results) was available.

Exclusion criteria were: (1) image quality did not meet diagnostic requirements (breast density classification $\geq C$ or blurred lesion borders); (2) history of anticancer treatment (chemotherapy/radiotherapy) or surgery in the area of the lesion within 6 months; (3) key variables were missing from clinical data ($>10\%$ missing).

To ensure the reliability of pathology, only paraffin results from surgical resection were used in this study, and negative results of frozen section pathology and puncture pathology were discarded. Finally, a total of 104 patients were enrolled, including 56 malignant cases and 48 benign cases. Lesions were visible in both positions in 98 cases, and only 6 cases showed lesions in one position. Patients' clinical information including age, menarche age, age of first delivery, lactation duration, family history of tumor, menopause, number of births, surgical history and clinical palpation were recorded.

Image Acquisition and BI-RADS Classification

Patients routinely underwent craniocaudal and mediolateral oblique projections using breast digital machines. Image quality and reporting were consistently assessed by two breast radiologists with more than 15 years of experience. The BI-RADS classification was performed through independent and blind review by three breast radiologists with ≥ 15 years of seniority based on 2019 WHO classification system, lesions identified as BI-RADS category 4A were enrolled for further analysis.

Lesion Segmentation

Lesion segmentation was performed independently by two radiologists blinding to patients' information and pathologic results using ITK-SNAP software (v3.6.0). Target lesion localisation in the baseline images was determined based on the inverse matching of the region of progression at the final examination. Each image corresponds to a unique lesion (no multiple images for the same lesion), and patients with multiple images ($n=98$) were included. A total of 202 ROIs (94 benign and 108 malignant) were outlined and randomly divided into a training set ($n=140$) and a test set ($n=62$) in a 7:3 ratio, with stratification to ensure a balanced ratio of benign to malignant. Given the retrospective data constraints where complete patient metadata was unavailable, all images were treated as independent samples during analysis. We note that patient-level isolation will be strictly enforced in clinical deployment. In clinical deployments, all data from the same patient will be forcibly segregated to a single application scenario (eg, training-only or prediction-only), ensuring safe model generalisation.

Reproducibility Analysis of Radiomics Features

To assess the interobserver reproducibility of extracted radiomics features, two radiologists performed the second extraction of radiomics features from 25 randomly chosen images after 1 month according to the same procedure. The intra-class correlation coefficient (ICC) was used to assess the intra-observer agreement, which was graded as very good when $ICC > 0.75$.¹²

Radiomics Signature Engineering

Extraction of Radiomics Features

In total, 1316 quantitative features were extracted based on the A.K. software (v3.0.1) covering morphological features ($n=14$) like volume, surface area, sphericity, etc.; first-order statistics features ($n=18$) including energy, entropy, skewness, kurtosis, etc.; textural features ($n=1284$) containing grey scale covariance matrix (GLCM), Gy-level run length matrix (GLRLM), Gy-level size zone matrix (GLSZM), Gy-level dependence matrix (GLDM) and neighbourhood grey scale difference matrix (NGTDM). Wavelet, square, gradient, square root, and exponential features were included. All features were normalised by Z-score to eliminate magnitude differences.

Feature Selection and Model Development

Spearman rank correlation coefficients were measured for any two features, and for the group of highly correlated features ($|\rho| > 0.9$), redundancy was removed by the following steps: a) Calculate univariate correlations (Spearman ρ) of all features with the target variable; b) Construct a hierarchical clustering tree (distance = $1 - |\rho|$, connectivity

= complete); c) Retain in each cluster the feature with the strongest correlation with the target variable and the highest clinical interpretability. A 10-fold cross-validation using the least absolute shrinkage and selection operator (LASSO) was used to determine the value of λ , retaining non-zero coefficient features. Features with $P < 0.1$ in univariate analyses were included in multivariate logistic regression analysis to identify the independent predictors.

Predictive Model Construction and Validation

A radiomics model was constructed based on logistic regression analysis for the radiomics features. Radiological scores (radscore) were generated based on the corresponding weighting coefficients in the radiological model, and the radscore was calculated for each case. The clinical-imaging fusion model was constructed by incorporating the radscore and independent clinical predictors (age, family history, etc.) into the multivariate model. Finally, column line plots were drawn using the R language rms package (v6.2–0), and internal validation was performed by the Bootstrap method (1000 resamples) to evaluate the model's differentiation, calibration, and clinical utility using a combination of area under the receiver operating characteristic (ROC) curve (AUC), calibration curve, and decision curve analysis (DCA). The calibration plot presented here serves primarily to evaluate internal consistency of predicted probabilities within cohort, rather than to assert generalizability.

Statistical Analysis

Statistical analysis was performed using SPSS software (version 22.0, IBM, USA). Normality of variances were evaluated through the one-sample Kolmogorov–Smirnov (K-S) test. Normally distributed continuous variables were demonstrated as mean \pm standard deviation and compared with Student's *t*-test, Mann–Whitney test was used for comparison of non-normally distributed continuous variables described as median and IQR. Chi-square test or Fisher's exact probability method was used for categorical variables reported as frequencies with percentages. ROC analysis was performed, the classification performance of models was evaluated by AUC and compared using DeLong test. Accuracy, sensitivity, specificity, and other indicators were calculated. All statistical tests were two-sided, and P -values < 0.05 were considered to be statistically significant.

Results

Clinical Characteristics and Pathology of the Patients

A total of 104 patients were enrolled in this study with a median age of 50 years (range 27–89 years). Malignant pathologies including intraductal carcinoma/lobular carcinoma in situ, non-specific invasive carcinoma, special types of carcinoma (tubular carcinoma, papillary carcinoma, mucinous carcinoma) and invasive lobular carcinoma, were found in 57 cases (57/104, 54.8%). Benign pathologies including breast adenopathy/hyperplasia, fibrocystic breast disease with atypical hyperplasia, fibroadenoma, phyllodes tumor, and sclerosing adenopathy were found in 47 cases (47/104, 45.2%). The specific pathological distribution results of the patients are shown in Table 1.

Univariate analysis was performed on the general clinical characteristics of the patients according to the pathological benign and malignant results of the patients. There was no significant difference between the benign group and the malignant group in terms of age, family history of tumor, menarche time, menopause, fetal history, primiparous age, breastfeeding history and gland type ($P > 0.05$). There were significant differences between the benign group and the malignant group in terms of surgical history ($P = 0.025$) and positive clinical palpation ($P = 0.026$). See Table 2 for details.

Table 1 Pathological Results of the 104 Patients

Pathology	Intraductal Carcinoma/ Lobular Carcinoma in situ	Invasive Carcinoma (Non-Specific)	Special Types of Carcinoma	Invasive Lobular Carcinoma	Breast Adenopathy/ Hyperplasia	Fibrocystic Breast Disease	Fibroadenoma	Phyllodes Tumor	Sclerosing Adenosis
Malignant	17	33	5	2	0	0	0	0	0
Benign	0	0	0	0	34	5	3	1	4

Table 2 General Characteristics of the 104 Patients

	Overall	Benign Tumor	Malignant Tumor	P-value
Age, years	50.0 (46.0–55.0)	50.0 (46.0–52.0)	50.0 (46.0–57.0)	0.461
Menarche age, years	14.0 (13.0–15.0)	14.0 (13.0–15.0)	13.0 (13.0–15.0)	0.221
Age of first delivery, years	27.0 (25.0–29.0)	27.0 (25.0–29.0)	27.0 (25.0–28.0)	0.824
Lactation duration, years	8.0 (2.5–12.0)	6.0 (1.5–10.0)	8.0 (4.0–12.0)	0.135
Family history of tumor, years	30 (28.85%)	14 (29.17%)	16 (28.57%)	0.809
Menopause, years	40 (38.46%)	18 (37.50%)	22 (39.29%)	1.000
Number of births				0.583
0	8 (7.69%)	4 (8.33%)	4 (7.14%)	
1	73 (70.19%)	35 (72.92%)	38 (67.86%)	
2	21 (20.19%)	9 (18.75%)	12 (21.43%)	
3	2 (1.92%)	0 (0.00%)	2 (3.57%)	
Surgical history, years	22 (21.15%)	5 (10.42%)	17 (30.36%)	0.025
Clinical palpation, positive	67 (64.42%)	25 (52.08%)	42 (75.00%)	0.026

Notes: missing value using multiple interpolation methods, the proportion of the missing value of each variable does not exceed 6%, continuity variables with median (interquartile spacing), difference between groups with Mann–Whitney *U*-test; classification variables using frequency (percentage), difference between groups by chi-square test, when the data is limited, Fisher exact probability test is used.

Image Annotation Reproducibility Validation

Both radiologists had better diagnostic results in terms of inter-observer agreement as well as intra-observer agreement (ICC of 0.882–0.986).

Predictive Feature Screening and Model Construction

Four key texture features were filtered by LASSO regression: lbp-3D-m2_grlm_RunLengthNonUniformityNormalized, wavelet-LLH_grlm_RunLengthNonUniformity, and original_grlm_Run-Entropy, and wavelet-LLH_gldm_SmallDependence LowGrayLevelEmphasis. According to the results of univariate and multivariate analyses for breast malignancy in the training group, the presence of surgical history and positive clinical palpation were both demonstrated to be independent predictors of breast malignancy (both $P < 0.05$).

Comparison of Model Performance

In this study, the performance of the radiomics model and the radiomics-clinical fusion model were compared in terms of accuracy, AUC values, sensitivity, specificity, positive predictive value and negative predictive value (as shown in Table 3).

Figure 1a and b showed the ROC curve and AUC value of the radiomics/fusion model in the training group and the testing group, respectively. There was no significant difference in the AUC of two models in training group ($P = 0.096$) and test group ($P = 0.494$). Figure 2a and b demonstrated the calibration curves of the models for the training and test

Table 3 Comparison of Performance of Radiomics Model and Radiomics-Clinical Fusion Model

Item	Train		Test	
	Radiomics-Only	Fusion	Radiomics-Only	Fusion
Accuracy	0.801	0.879	0.803	0.883
AUC	0.858	0.923	0.860	0.904
Sensitivity	0.871	0.800	0.833	0.861
Specificity	0.714	0.957	0.774	0.917
Positive prediction	0.791	0.949	0.781	0.939
Negative prediction	0.818	0.827	0.828	0.815

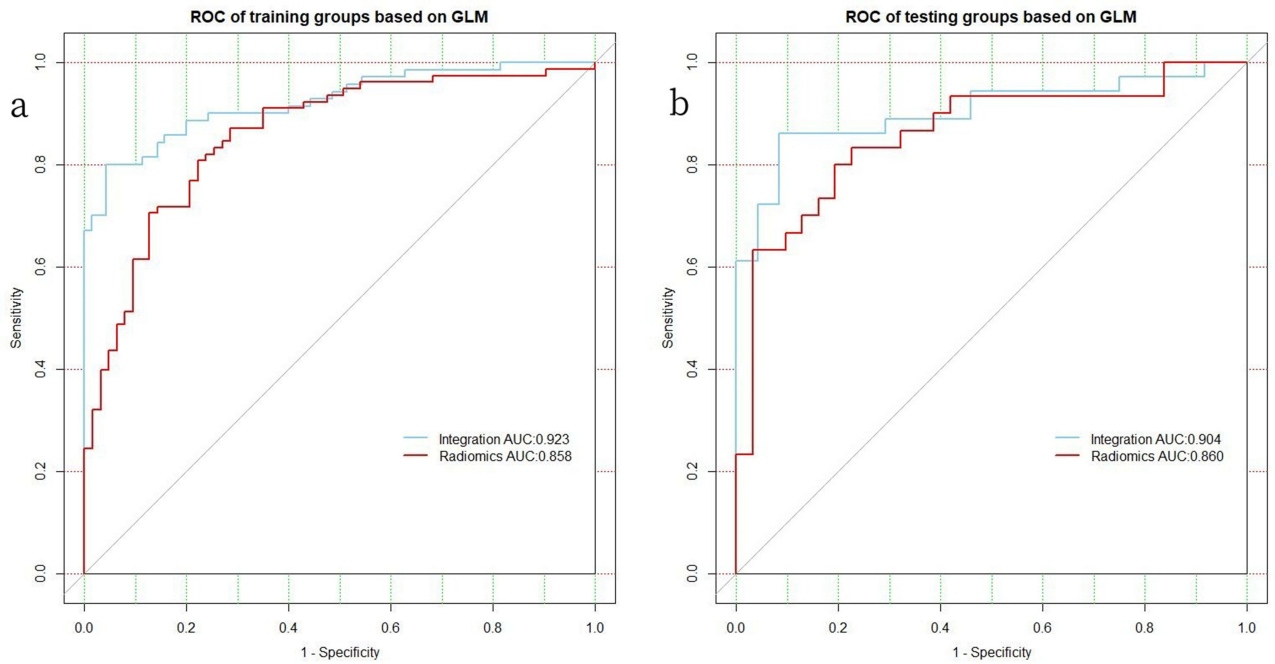


Figure 1 ROC curves of radiomics model and radiomics-clinical fusion model. (a) The training set; (b) The testing set.

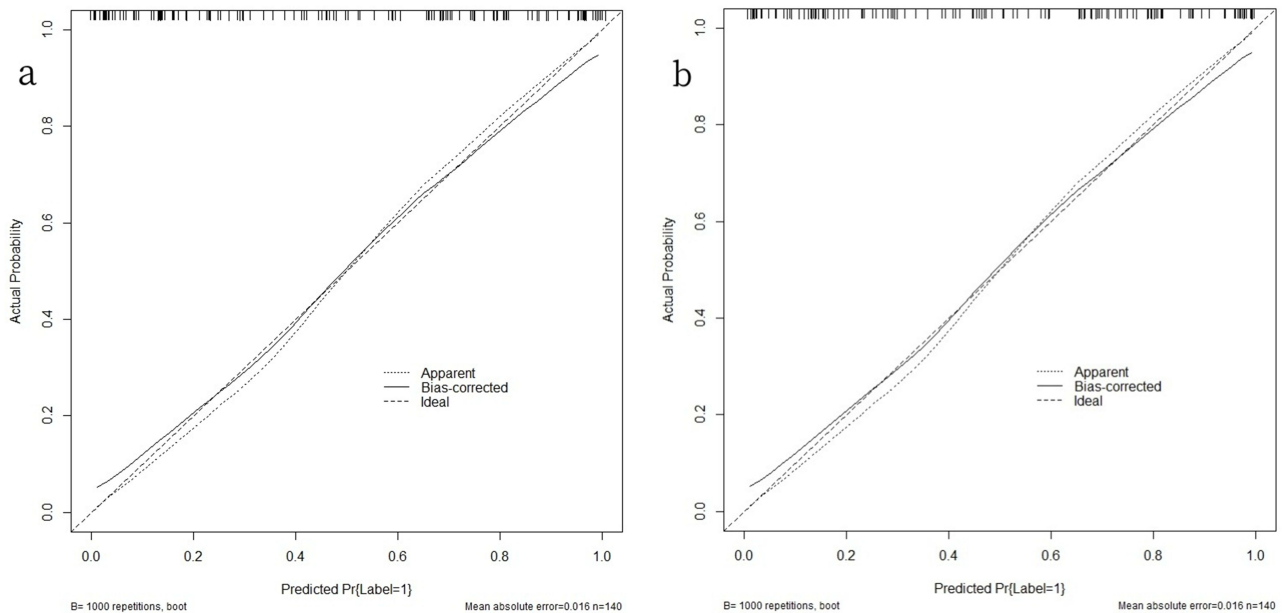


Figure 2 The calibration curves of radiomics model and radiomics-clinical fusion model. (a) The training set; (b) The testing set.

groups, respectively, and the diagonal dashed and solid lines match perfectly, indicating good internal consistency. The calibration plot suggests good agreement between predicted and observed probabilities in this dataset, though this may reflect intrinsic cohort characteristics rather than model generalizability. Figure 3a and b exhibited the DCA decision curves of the fused model in the training set and the test set, respectively. When the risk thresholds are 10%–95% in the training and 20–95% in the validation group, the fusion model demonstrated good clinical usefulness.

The nomogram of the fusion model was shown in Figure 4. The importance and effectiveness of the rad-score was outstanding compared to other clinical features.

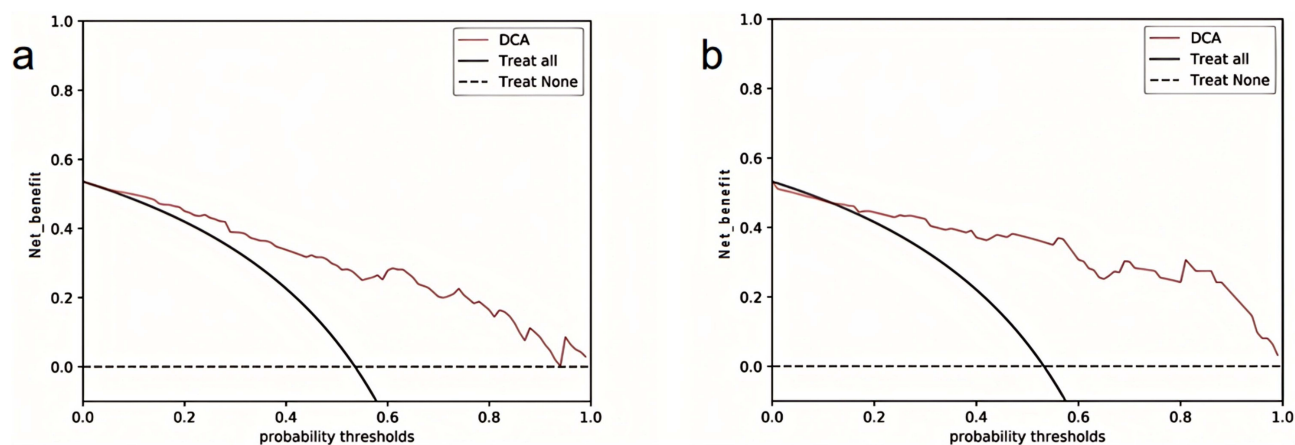


Figure 3 The DCA decision curves of radiomics-clinical fusion model. (a) The training set; (b) The testing set.

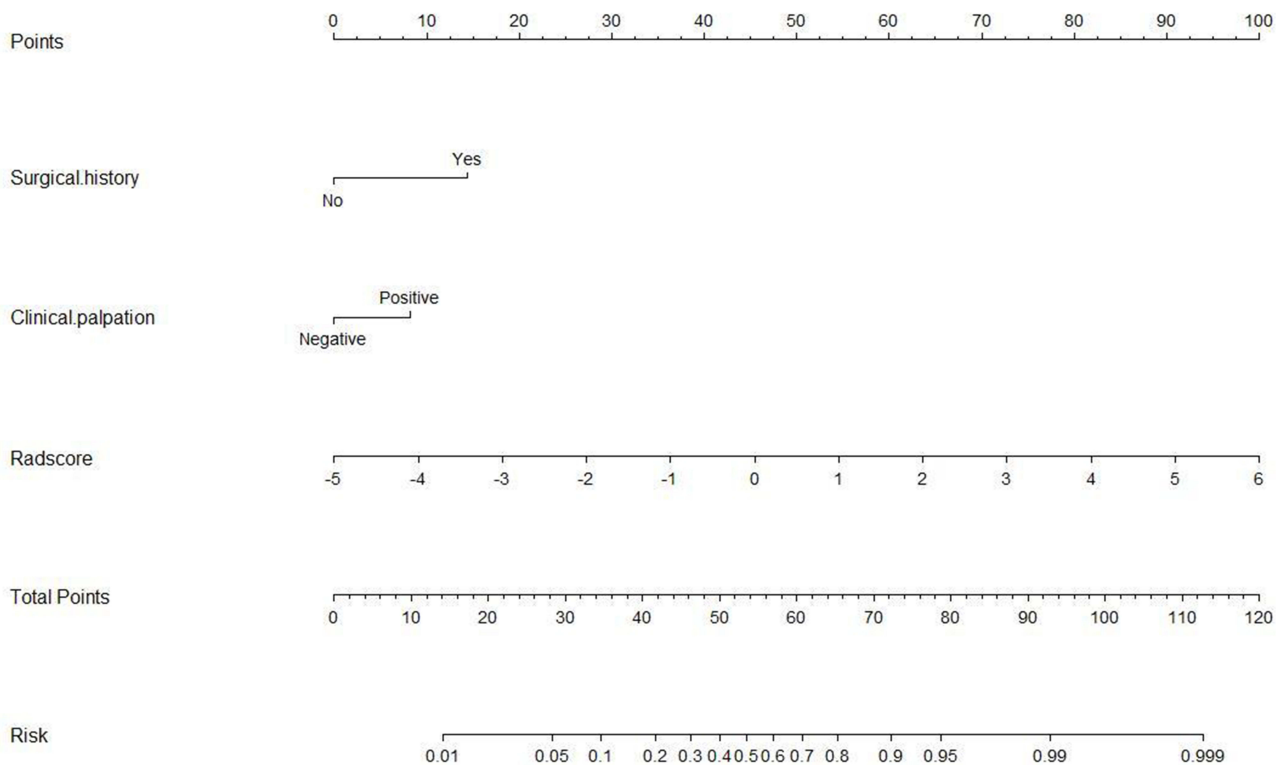


Figure 4 The nomogram of the fusion model.

Discussion

In this study, we developed a fusion Nomogram model based on radiomics features and clinical characteristics to predict the benign and malignant progression of BI-RADS 4A category lesions. The model included not only radiomics features but also integrates clinical features such as palpation positivity and history of breast surgery to enhance the accuracy of predictions. Our results indicate that the AUC value of the fusion model (0.904) was higher than that of the radiomics-only model (0.860). Although no statistically significant difference was found between the two models, the fusion model demonstrated superior predictive performance in both the training and testing cohorts. The innovation of this study lies in

its focus not only on the differentiation between benign and malignant 4A lesions but also on the prediction of the progression trend of these lesions during follow-up, which is of significant importance for clinical decision-making.

This study has several highlights. While most existing studies concentrate on the differentiation of benign and malignant 4A lesions, this study further predicts the progression trend of these lesions during follow-up, which is of great value in reducing unnecessary biopsies and surgeries. Furthermore, mammography is cost-effective compared to MRI examinations which require high costs and may pose the risk of allergies. The predictive model based on the radiomics features of mammographic imaging made it more suitable for primary hospitals and large-scale screening.¹⁴

The four radiomics features selected in this study all belong to texture features, which can reflect the roughness and smoothness of the lesion area, indirectly reflecting the different genetic, molecular, and protein compositions of the lesion area. The early manifestations and stroma of benign and malignant lesions are completely different, and the stroma of different types of malignant lesions can also be completely different. Therefore, the four radiomics features extracted in this study can perform well in predicting the benign and malignant nature of 4A category lesions. Similar to previous studies (10.1016/j.ejrad.2019.108711; 10.1007/s00330-023-09513-3) for differentiating between benign and malignant breast lesions initially categorized as BI-RADS 4A, our study confirmed that texture features still hold significant value in predicting the benign and malignant outcomes of BI-RADS 4A lesions, demonstrating its excellent ability to mine deep information for reflection the internal heterogeneity of breast lesions. Meanwhile, Zhang et al (10.1007/s00330-023-09513-3) extracted more comprehensive radiomic features from both the lesion and the surrounding area. The radiomic features from the surrounding area of the lesion were ultimately incorporated into their nomogram, which confirmed the value of the surrounding tissue and microenvironment in differentiating the benign and malignant nature of breast lesions, which provided a direction that needs to be further supplemented and improved in our future research. As demonstrated by Li et al,¹⁵ radiomics has gained widespread recognition for its ability to differentiate between benign and malignant lesions and various types of breast cancer.

Three clinical characteristics were selected in this case series. Palpation plays the most significant role in differentiating benign and malignant lesions. A negative palpation often suggests a benign lesion, while a palpable thickening or mass in the palpated lesion area may suggest a malignant lesion. Almost all relevant guidelines recommend physical examination for patients.¹⁶ It is worth noting that although a positive clinical examination significantly improves the predictive efficacy of breast cancer, one should be wary of the diagnostic complexity of high-grade (grade 3) tumours due to their rapid growth. The literature suggests that such tumours may overlap with benign lesions due to atypical imaging features (eg, blurred borders, absence of calcification, or heterogeneous enhancement),¹⁷ and that their lack of typical palpation features may further increase the risk of misdiagnosis in the first instance.¹⁸ Although the present study was limited by retrospective data and was unable to directly quantify the misdiagnosis rate of high-grade tumours, in light of the available evidence, we recommend the use of multimodal assessment (eg, dynamic enhanced MRI combined with molecular marker testing) for rapidly progressive lesions in clinical practice to compensate for the limitations of a single examination. Future prospective studies will focus on validating the efficacy of fusion models in such high-risk subgroups. A history of breast surgery was also found to be significantly different in benign and malignant lesions in this study, and literature reports that a history of benign or malignant breast tumor surgery can increase the risk of breast cancer on the same side or the contralateral side, and has become one of the predictive factors in some risk prediction models.^{19–22} The protective effect of breastfeeding on breast cancer is well known, and previous studies have proven that breastfeeding can reduce the risk of luminal and triple-negative breast cancer, but it cannot reduce the risk of HER2-positive breast cancer.^{23–26} These findings emphasize the importance of considering multiple factors in the risk assessment of breast cancer. Although the clinical-radiomics fusion model developed in this study significantly improves predictive efficacy, it is important to emphasise that a single examination tool (either clinical palpation or imaging model) cannot replace a comprehensive diagnostic process. According to the European guidelines for the management of breast cancer,²⁷ all imaging abnormalities or clinically suspicious lesions need to be confirmed by triple assessment, especially for lesions with ambiguous borders or rapid progression (eg, high-grade tumours). The radiomics model of this study can be used as a complementary tool for imaging assessment to assist in the identification of atypical malignant lesions (eg, heterogeneous enhancement or absence of microcalcification) in conventional BI-RADS classification, but its predictive results still need to be interpreted in conjunction with puncture biopsy or intraoperative frozen pathology.²⁸ Future work

will further integrate intraoperative pathological features (eg, Ki-67 index or molecular typing) to construct a multimodal decision-making system covering the entire diagnostic process to guide individualised treatment more precisely.

Despite the achievements of our study, there were some limitations. First, this study is a single-center, retrospective study with a limited sample size, which may lead to selection bias. Subsequent studies should expand the sample size and be validated in multi-center, prospective studies. Our calibration analysis is limited by the use of a single-institution dataset. While this helps assess internal consistency, it does not account for inter-institutional variability in imaging protocols or patient demographics.²⁹ We emphasized that our model's clinical applicability requires further external testing. Second, this study only included two-dimensional images of mammography and failed to include tomographic images or other imaging data, limiting the scope of applicability of the results. Future studies will include multimodal medical imaging to broaden the application range of the model.

In summary, the fusion Nomogram model developed in this study based on radiomics and clinical characteristics showed good potential in predicting the benign and malignant progression of BI-RADS 4A category lesions. This tool is expected to provide more accurate assistance in the clinical risk assessment of 4A category lesions, thereby reducing unnecessary surgeries and improving the quality of life for patients.

Data Sharing Statement

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Ethics Approval and Consent to Participate

This study was approved by the Ethics Committee of The Second Affiliated Hospital of Guangzhou University of Chinese Medicine (approval number: ZE2021-111). Informed consent was obtained from all the participants. All methods were carried out in accordance with Declaration of Helsinki.

Consent for Publication

All authors confirmed that written informed consent was obtained from all subjects and/or their legal guardian(s).

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

The work was not funded by any funding.

Disclosure

Guoyan Yao, Lijun Chen and Tingfan Wu They are the co-first authors for this study. The authors declared that they have no conflicts of interest regarding this work.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74(3):229–263. doi:10.3322/caac.21834
2. Huang J, Chan PS, Lok V, et al. Global incidence and mortality of breast cancer: a trend analysis. *Aging.* 2021;13(4):5748–5803. doi:10.18632/aging.202502
3. D'Orsi CJ, Sickles EA, Mendelson EB, et al. ACR BI-RADS® Atlas. Breast imaging reporting and data system. Reston: American College of Radiology; 2013.
4. Elezaby M, Li G, Bhargavan-Chatfield M, et al. ACR BI-RADS assessment category 4 subdivisions in diagnostic mammography: utilization and outcomes in the national mammography database. *Radiology.* 2018;287(2):416–422. doi:10.1148/radiol.2017170770

5. Wang S, Sun Y, Li R, et al. Diagnostic performance of perilesional radiomics analysis of contrast-enhanced mammography for the differentiation of benign and malignant breast lesions. *Eur Radiol.* 2022;32(1):639–649. doi:10.1007/s00330-021-08134-y
6. Glechner A, Wagner G, Mitus JW, et al. Mammography in combination with breast ultrasonography versus mammography for breast cancer screening in women at average risk. *Cochrane. Database Syst Rev.* 2023;3(3):CD009632.
7. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun.* 2020;11(1):1236. doi:10.1038/s41467-020-15027-z
8. Mao N, Shi Y, Lian C, et al. Intratumoral and peritumoral radiomics for preoperative prediction of neoadjuvant chemotherapy effect in breast cancer based on contrast-enhanced spectral mammography. *Eur Radiol.* 2022;32(5):3207–3219. doi:10.1007/s00330-021-08414-7
9. Jiang M, Li CL, Luo XM, et al. Ultrasound-based deep learning radiomics in the assessment of pathological complete response to neoadjuvant chemotherapy in locally advanced breast cancer. *Eur J Cancer.* 2021;147:95–105. doi:10.1016/j.ejca.2021.01.028
10. Huang Y, Wei L, Hu Y, et al. Multi-parametric MRI-based radiomics models for predicting molecular subtype and androgen receptor expression in breast cancer. *Front Oncol.* 2021;11:706733. doi:10.3389/fonc.2021.706733
11. Lei C, Wei W, Liu Z, et al. Mammography-based radiomic analysis for predicting benign BI-RADS category 4 calcifications. *Eur J Radiol.* 2019;121:108711. doi:10.1016/j.ejrad.2019.108711
12. Wengert GJ, Pipan F, Almohanna J, et al. Impact of the Kaiser score on clinical decision-making in BI-RADS 4 mammographic calcifications examined with breast MRI. *Eur Radiol.* 2020;30(3):1451–1459. doi:10.1007/s00330-019-06444-w
13. Hao W, Gong J, Wang S, et al. Application of MRI radiomics-based machine learning model to improve contralateral BI-RADS 4 lesion assessment. *Front Oncol.* 2020;10:531476. doi:10.3389/fonc.2020.531476
14. Bickelhaupt S, Paech D, Laun FB, et al. Maximum intensity breast diffusion MRI for BI-RADS 4 lesions detected on X-ray mammography. *Clin Radiol.* 2017;72(10):900.e1–900.e8. doi:10.1016/j.crad.2017.05.017
15. Li H, Mendel KR, Lan L, et al. Digital mammography in breast cancer: additive value of radiomics of breast parenchyma. *Radiology.* 2019;291(1):15–20. doi:10.1148/radiol.2019182310
16. Gradishar WJ, Moran MS, Abraham J, et al. NCCN guidelines insights: breast cancer, version 4.2021. *J Natl Compr Canc Netw.* 2021;19(5):484–493. doi:10.6004/jnccn.2021.0023
17. Dahlblom V, Dustler M, Tingberg A, Zackrisson S. Breast cancer screening with digital breast tomosynthesis: comparison of different reading strategies implementing artificial intelligence. *Eur Radiol.* 2023;33(5):3754–3765. doi:10.1007/s00330-022-09316-y
18. Buchberger W, Niehoff A, Obrist P, DeKoekkoek-Doll P, Dünser M. Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. *Semin Ultrasound CT MR.* 2000;21(4):325–336. doi:10.1016/S0887-2171(00)90027-1
19. Yu ZG, Jia CX, Geng CZ, et al. Risk factors related to female breast cancer in regions of Northeast China: a 1:3 matched case-control population-based study. *Chin Med J.* 2012;125(5):733–740.
20. Ward EM, DeSantis CE, Lin CC, et al. Cancer statistics: breast cancer in situ. *CA Cancer J Clin.* 2015;65(6):481–495. doi:10.3322/caac.21321
21. Chuba PJ, Hamre MR, Yap J, et al. Bilateral risk for subsequent breast cancer after lobular carcinoma-in-situ: analysis of surveillance, epidemiology, and end results data. *J Clin Oncol.* 2005;23(24):5534–5541. doi:10.1200/JCO.2005.04.038
22. Lambertini M, Santoro L, Del Mastro L, et al. Reproductive behaviors and risk of developing breast cancer according to tumor subtype: a systematic review and meta-analysis of epidemiological studies. *Cancer Treat Rev.* 2016;49:65–76. doi:10.1016/j.ctrv.2016.07.006
23. Giudici F, Scaggiante B, Scomersi S, et al. Breastfeeding: a reproductive factor able to reduce the risk of luminal B breast cancer in premenopausal white women. *Eur J Cancer Prev.* 2017;26(3):217–224. doi:10.1097/CEJ.0000000000000220
24. Sisti JS, Collins LC, Beck AH, et al. Reproductive risk factors in relation to molecular subtypes of breast cancer: results from the nurses' health studies. *Int J Cancer.* 2016;138(10):2346–2356. doi:10.1002/ijc.29968
25. Kotsopoulos J, Lubinski J, Salmena L, et al. Breastfeeding and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res.* 2012;14(2):R42. doi:10.1186/bcr3138
26. Mao N, Shi Y, Lian C, et al. Intratumoral and peritumoral radiomics for preoperative prediction of neoadjuvant chemotherapy effect in breast cancer based on contrast-enhanced spectral mammography. *Eur Radiol.* 2022;32(1):639–649.
27. Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up† [published correction appears in *Ann Oncol.* 2019 Oct 1;30(10):1674. doi: 10.1093/annonc/mdz189.] [published correction appears in *Ann Oncol.* 2021 Feb;32(2):284. doi: 10.1016/j.annonc.2020.08.2158.]. *Ann Oncol.* 2019;30(8):1194–1220. doi:10.1093/annonc/mdz173
28. American College of Radiology. ACR practice guideline for the performance of contrast-enhanced magnetic resonance imaging (MRI) of the breast. 2023. Available from: <https://www.acr.org/Clinical-Resources/Practice-Parameters>. Accessed September 11, 2025.
29. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21(1):128–138. doi:10.1097/EDE.0b013e3181c30fb2

International Journal of Women's Health

Publish your work in this journal

The International Journal of Women's Health is an international, peer-reviewed open-access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of women's healthcare including gynecology, obstetrics, and breast cancer. 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/international-journal-of-womens-health-journal>

Dovepress
Taylor & Francis Group