



# Volumetric phase-specific three-dimensional radiomics and machine learning for differentiation of non-mass enhancement in breast magnetic resonance imaging

Serhat Akış<sup>1</sup>  
 Pınar Balcı<sup>2</sup>  
 Muhammet Baldan<sup>3</sup>  
 Melike Zeynep Akış<sup>2</sup>  
 Işıl Başara Akın<sup>2</sup>  
 Süleyman Aksoy<sup>4</sup>  
 Ali İbrahim Sevinç<sup>4</sup>  
 Merih Güray Durak<sup>5</sup>

<sup>1</sup>University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Clinic of Radiology, İzmir, Türkiye

<sup>2</sup>Dokuz Eylül University Faculty of Medicine, Department of Radiology, İzmir, Türkiye

<sup>3</sup>Acıbadem University, Faculty of Medicine, Department of Bioinformatics and Biostatistics, İstanbul, Türkiye

<sup>4</sup>Dokuz Eylül University Faculty of Medicine, Department of General Surgery, İzmir, Türkiye

<sup>5</sup>Dokuz Eylül University Faculty of Medicine, Department of Pathology, İzmir, Türkiye

**Corresponding author:** Serhat Akış

**E-mail:** serhatakis@gmail.com

Received 27 January 2026; revision requested 01 March 2026; last revision requested 08 April 2026; accepted 13 April 2026.



Epub: 28.04.2026

Publication date:

DOI: 10.4274/dir.2026.263899

## PURPOSE

Non-mass enhancement (NME) in breast magnetic resonance imaging (MRI) is a diagnostically challenging entity due to overlapping benign and malignant features, observer variability, and high false-positive rates. This study evaluated the diagnostic performance of three-dimensional (3D) volumetric radiomics and multiple machine learning (ML) algorithms, using early (2<sup>nd</sup>) and late (7<sup>th</sup>) post-contrast phases to differentiate benign from malignant NMEs.

## METHODS

A total of 110 NMEs (86 benign and 24 malignant) from 108 patients were analyzed. Radiological features were recorded. Radiomics features were extracted from manual 3D segmentations using LIFEx software. Multivariate logistic regression (LR) and supervised ML algorithms—LR, support vector machine, random forest, and gradient boosting—were applied. The methodological quality was assessed using the Multicenter Evaluation of Radiomics in Clinical Studies framework.

## RESULTS

A total of 54 lesions were histopathologically confirmed, and 56 were confirmed by follow-up. Among the 56 lesions evaluated by follow-up, 34 remained stable ( $\geq 24$  months), whereas 22 showed regression ( $\geq 6$  months). Distribution, internal enhancement, size, and laterality differed significantly between benign and malignant NMEs ( $P < 0.05$ ). Radiomics analysis extracted 123 features, of which 92 on the early and 80 on the late post-contrast images were significant for benign–malignant differentiation ( $P < 0.05$ ). In the early phase, combining all radiomics features increased specificity from 78% to 98% and accuracy from 82% to 93%. ML models further improved performance, achieving specificity up to 99% and area under the curve (AUC) values exceeding 0.91. Similar improvements were observed on the late phase, with accuracies up to 91% and AUC values up to 0.93.

## CONCLUSION

Volumetric 3D radiomics, combined with ML, using early and late post-contrast phases improves diagnostic accuracy and specificity for NME on breast MRI.

## CLINICAL SIGNIFICANCE

Integrating 3D radiomics and ML into breast MRI evaluation supports more accurate decision-making in NME and may reduce unnecessary biopsies.

## KEYWORDS

Breast, machine learning, magnetic resonance imaging, non-mass enhancement, radiomics

According to the Breast Imaging-Reporting and Data System (BI-RADS) Atlas,<sup>1</sup> non-mass enhancement (NME) in breast magnetic resonance imaging (MRI) refers to abnormal areas of enhancement that do not fulfill the criteria of a mass and are not space-occupying. Despite the high sensitivity of MRI in breast diseases,<sup>2</sup> NME is the leading cause of false-positive findings in breast MRI.<sup>3</sup> The evaluation of NME remains a diagnostic challenge due to the considerable intra- and inter-observer variability in its interpretation<sup>4,5</sup> and the substantial overlap between benign and malignant findings.<sup>6</sup>

There is also a lack of consistency in the literature regarding the relationship between the distribution patterns or internal enhancement characteristics of NME and the likelihood of malignancy. Although some studies have suggested correlations between specific enhancement patterns and malignant outcomes, others have reported conflicting results.<sup>7-14</sup> These discrepancies further complicate the diagnostic process and limit reproducibility.

In recent years, computer-aided diagnostic approaches have gained growing attention as a potential solution in radiological decision-making. Radiomics and texture analysis allow for the quantitative extraction of imaging features that are not discernible to the human eye, thereby providing a more objective assessment of lesion characteristics.<sup>15-17</sup> Machine learning (ML) techniques, by leveraging large amounts of imaging and clinical data, have shown promising results in differentiating benign from malignant breast lesions.<sup>18</sup>

The integration of computational methods into breast MRI has the potential to improve diagnostic accuracy, reduce variability among radiologists, and minimize unnecessary biopsies. Previous radiomics, ML, and deep learning studies on NME have predominantly relied on two-dimensional (2D), slice-based analyses, focusing on a single or limited number of representative images rather than the entire lesion volume.<sup>19-22</sup> To our knowledge, no previous study has comprehensively applied three-dimensional (3D) radiomics and ML for the characterization of NME.

Therefore, our study aims to evaluate the diagnostic performance of 3D radiomics-based texture analysis and ML models in distinguishing benign from malignant NME in breast MRI. This study aims to provide more objective and reproducible diagnostic tools to support clinical decision-making.

## Methods

### Study design and population

The study was approved by the Dokuz Eylül University Non-invasive Clinical Research Ethics Committee (file number: 8842-GOA; date: 03.04.2024). Written informed consent was obtained from all patients before breast MRI examination as part of routine clinical practice; however, due to the retrospective design of the study, separate informed consent for study participation was waived. This retrospective cross-sectional study included 110 NMEs in 108 patients who underwent dynamic contrast-enhanced breast MRI

(DCE-MRI) between May 2015 and August 2024. During the study period, 12,346 breast MRI examinations were performed at the study institution. Cases were screened for the presence of NME lesions. The inclusion criteria were as follows: age  $\geq 18$  years, DCE-MRI with adequate imaging parameters, and NME lesions with either histopathological confirmation or imaging follow-up. Exclusion criteria included inadequate MRI parameters, poor image quality or artifacts, enhancement smaller than 5 mm, and the absence of a definitive diagnosis due to a lack of biopsy or insufficient follow-up. Lesions showing regression were included if follow-up was  $\geq 6$  months, whereas lesions classified as stable were included only if follow-up was  $\geq 24$  months. After applying these criteria, 110 NMEs in 108 patients were included in the final analysis (Figure 1).

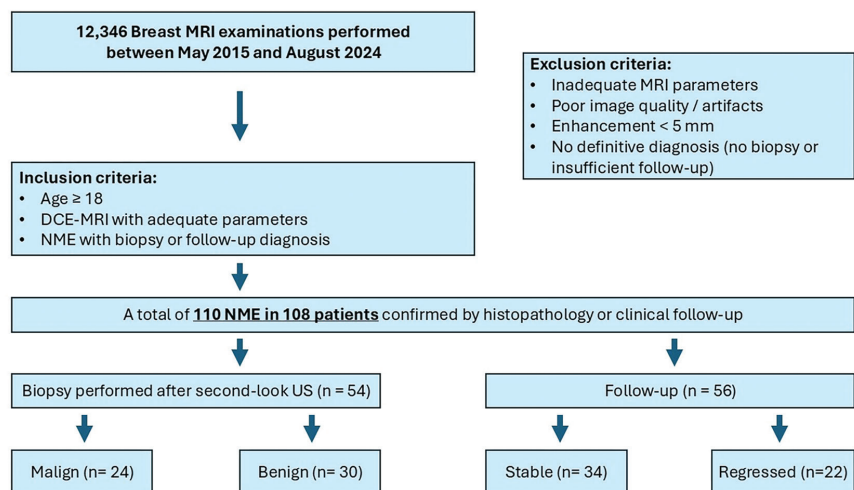
Of these lesions, 54 underwent biopsy after second-look ultrasound (SLU), yielding 24 malignant and 30 benign lesions. The remaining 56 lesions were evaluated by imaging follow-up, with 34 remaining stable and 22 showing regression during follow-up.

### Magnetic resonance imaging protocol and radiological assessment

All examinations were performed on a 1.5-T scanner (Philips Achieva Release 1.8, Eindhoven, The Netherlands) using a dedicated breast coil in the prone position. The protocol was the same for all patients: T1- and T2-weighted (W) sequences, diffusion-weighted imaging (DWI), and DCE 3D T1W gradient-echo imaging with and

**Main points**

- Non-mass enhancement (NME) on breast magnetic resonance imaging has substantial benign-malignant overlap and high false-positive rates, limiting diagnostic confidence with conventional radiological assessment alone.
- Three-dimensional volumetric radiomics extracted from both early (2<sup>nd</sup>) and late (7<sup>th</sup>) post-contrast phases significantly improved diagnostic specificity and overall accuracy for differentiating benign from malignant NMEs.
- Machine learning models integrating radiological features with recursive feature elimination-selected radiomics features achieved high performance (area under the curve > 0.91) and markedly increased specificity (up to 99%), suggesting potential to reduce unnecessary biopsies.



**Figure 1.** Flowchart of patient selection and study cohort. Between May 2015 and August 2024, a total of 12,346 dynamic DCE-MRI examinations were reviewed. After applying the inclusion and exclusion criteria, 110 non-mass enhancements from 108 patients were included. Of these, 54 lesions underwent biopsy after second-look ultrasound (24 malignant and 30 benign), whereas 56 lesions were managed with radiological follow-up (34 stable and 22 regressed). DCE, dynamic contrast-enhanced breast MRI; MRI, magnetic resonance imaging; NME, non-mass enhancement; US, ultrasound.

without fat-suppressed images. Gadolinium-based contrast agents (gadoterate or gadobutrol, 0.1 mmol/kg) were administered intravenously at 2 mL/s. Seven post-contrast subtraction series were generated. Lesion characteristics—distribution, internal enhancement, laterality, T2 signal intensity, maximum size, breast density (BD), and background parenchymal enhancement (BPE)—were evaluated by a dedicated breast radiologist with > 25 years of experience. In cases where MRI detected suspicious lesions and biopsy was recommended, SLU was performed for lesion correlation and biopsy planning. Representative examples are shown in Figure 2.

### Radiomics analysis

The 2<sup>nd</sup> and 7<sup>th</sup> post-contrast subtraction images were analyzed using the LIFEx software. Fat-suppressed sequences were preferred because the suppression of background fat enhances contrast uptake, making NME more conspicuous. Image preprocessing included voxel resampling, intensity discretization, and normalization to reduce variability related to image acquisition. Manual 3D segmentation of NMEs was performed by a radiologist with 5 years of experience, and segmentation examples are shown in Figure 3. A total of 123 radiomics features (morphologic, intensity, and second-order) were extracted. To ensure reproducibility, a random subset of 20 cases was independently segmented by a second radiologist (3 years of experience), with good interobserver agreement (Dice  $\approx$  0.80, intra-class correlation coefficient > 0.85,  $\kappa$  > 0.70).

### Machine learning models

Radiomics features were used both alone and in combination with radiological parameters to develop classification models. Initially, logistic regression (LR) was applied using radiological features alone. Subsequently, separate LR models were constructed using each radiomics feature group individually (morphological, intensity-based, and second-order texture features). Recursive feature elimination (RFE) was then performed to identify the 14 most relevant radiomics features from the 2<sup>nd</sup> and 7<sup>th</sup> post-contrast images. Using these selected features, an additional radiomics-based diagnostic model was generated, followed by a combined model integrating the 14 RFE-selected radiomics features with radiological parameters. In total, six diagnostic models were constructed for each contrast phase.

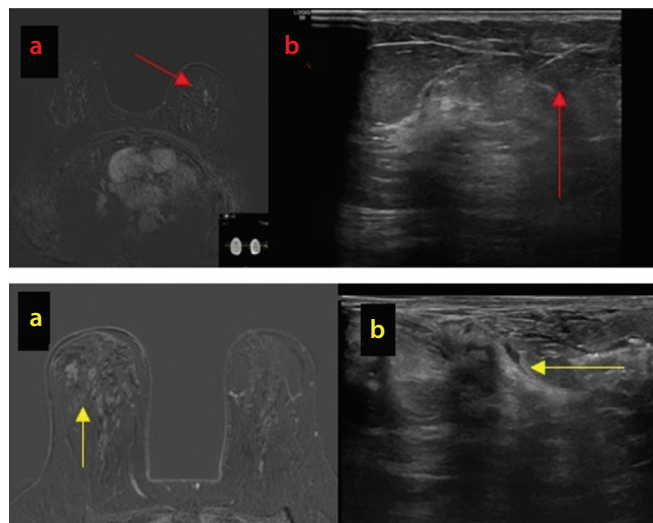
Support vector machine (SVM), random forest, and gradient boosting (GB) algorithms were tested using the combined model incorporating radiological features and the 14 selected radiomics features. All analyses were performed separately for the 2<sup>nd</sup> and 7<sup>th</sup> post-contrast images. Accuracy, sensitivity, specificity, and area under the curve (AUC) were reported. The overall study workflow is summarized in Figure 4.

Internal model evaluation was performed using 5-fold cross-validation. The dataset was divided into five folds, with four folds used for training and one fold used for val-

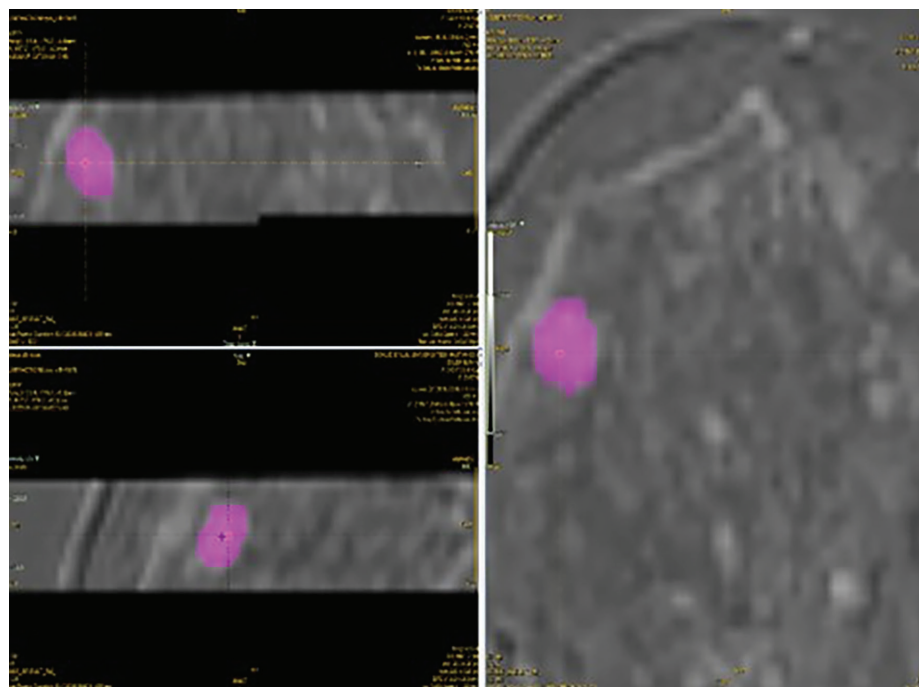
idation in each iteration. This process was repeated five times so that each fold served once as the validation set. RFE was applied before the cross-validation step to identify the 14 most relevant radiomics features for each phase. These 14 features are presented in Supplementary Tables 1 and 2.

### Statistical analysis

Statistical analyses were performed using chi-square tests for categorical variables and independent samples t-tests for continuous radiomics parameters. Multivariate LR was applied to evaluate the predictive



**Figure 2.** Second-look ultrasound findings of non-mass enhancements (NMEs). Representative cases demonstrate targeted second-look ultrasound correlation of breast MRI-detected NMEs: (a) linear-homogeneous enhancement, (b) focal clustered-ring enhancement. Lesions were successfully sampled using ultrasound-guided tru-cut biopsy. MRI, magnetic resonance imaging.



**Figure 3.** Segmentation workflow using radiomics software. Lesion segmentation was performed in three orthogonal planes using LIFEx. A total of 123 radiomics features were extracted.

value of radiological and radiomics features for malignancy. In total, five LR models were constructed: a model including radiological parameters alone, a general radiomics model, and three separate radiomics-based models incorporating morphological, histogram-based, and second-order texture features. To avoid multicollinearity, variance inflation factor values were assessed, and highly correlated variables were excluded. Model performance was assessed with receiver operating characteristic analysis, and odds ratios with 95% confidence intervals were reported.

According to the Multicenter Evaluation of Radiomics In Clinical Studies criteria,<sup>23</sup> the study demonstrated high methodological quality (overall score: 79.5%) (Appendix 1).

## Results

### Clinical and radiological findings

A total of 110 NMEs from 108 patients were included. Of these, 54 were confirmed histopathologically (24 malignant and 30 benign) and 56 by follow-up. Among the 56 follow-up lesions, 34 remained stable and had more than 24 months of follow-up (mean: 25 months), whereas 22 showed regression or resolution and had at least 6 months of follow-up (mean: 12.5 months). The mean age was  $47.9 \pm 9.7$  years in the benign group

and  $49.5 \pm 11.6$  years in the malignant group, with no significant difference between groups ( $P = 0.699$ ). Distribution and internal enhancement features demonstrated significant differences between groups (Table 1). Focal distribution occurred exclusively in benign NMEs (41.9%), whereas segmental distribution (54.2%) and heterogeneous enhancement (50%) predominated in malignant cases. Malignant NMEs were larger and more often right-sided ( $P < 0.05$ ). No significant differences were observed for BPE or BD between benign and malignant groups.

### Histopathological findings

Histopathological analysis was available for 54 lesions. Of these, 28 lesions (51.8%) were classified as benign without atypia, 2 lesions (3.7%) as benign with atypia, and 24 lesions (44.4%) as malignant. The most frequent benign histopathological findings were normal breast tissue, ductal epithelial hyperplasia, and fibroadenomatous changes (each 9.2%). The malignant group consisted predominantly of invasive carcinoma (21/24, 87.5%), with ductal carcinoma *in situ* (DCIS) accounting for the remaining cases (3/24, 12.5%).

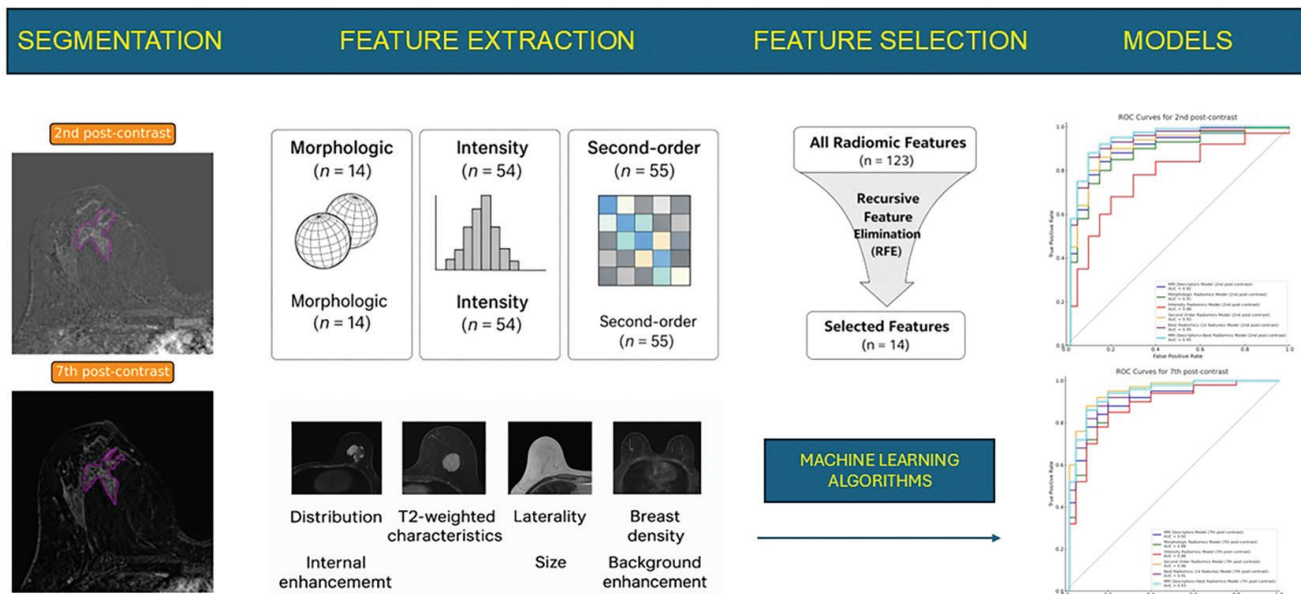
### Radiomics and logistic regression

Radiomics analysis extracted a total of 123 features, including 14 morphological, 54 intensity-based, and 55 second-order texture

features. A substantial proportion of the extracted features were significantly associated with benign–malignant differentiation. Specifically, 92 radiomics features were statistically significant in the early phase, whereas 80 features reached significance in the late phase (Table 2). Supplementary Materials 1 and 2 illustrate the radiomics features that demonstrated statistically significant associations for each contrast phase. The univariate analyses of radiomics features were performed as an exploratory step, and no formal correction for multiple comparisons was applied.

### Performance comparison on the 2<sup>nd</sup> post-contrast images

On multivariate LR analysis, radiological features alone demonstrated relatively low specificity. The inclusion of radiomics features substantially improved specificity. The most striking improvement was observed when all three radiomics feature groups—morphological, intensity-based, and second-order texture features—were combined, with specificity increasing from 78% to 98% and overall diagnostic accuracy improving from 82% to 93%. Among the individual radiomics feature groups, second-order texture features yielded the best model performance, achieving an AUC of 0.96.



**Figure 4.** Study of the radiomics and machine learning analysis. Following manual three-dimensional segmentation of non-mass enhancement lesions on dynamic contrast-enhanced breast MRI, radiomics features were extracted separately from the early (2<sup>nd</sup>) and late (7<sup>th</sup>) post-contrast phases. A total of 123 radiomics features, including morphological, intensity-based, and second-order texture features, were obtained. In addition, conventional radiological features (distribution, internal enhancement pattern, lesion size, laterality, T2-weighted signal characteristics, breast density, and background parenchymal enhancement) and age (not shown in the figure) were recorded. Feature selection was performed using recursive feature elimination, resulting in the selection of the 14 most relevant radiomics features. Classification models were subsequently developed using logistic regression and machine learning algorithms, and model performance was evaluated using ROC curve analysis for both post-contrast phases. MRI, magnetic resonance imaging; ROC, receiver operating characteristic.

When ML algorithms were applied independently, models trained solely on radiological features achieved a specificity of up to 96% and an accuracy of 88%. Incorporation of radiomics features into ML-based LR models resulted in further increases in diagnostic accuracy. The highest overall performance was obtained with the combined model integrating radiological features and 14 radiomics features selected using RFE.

When ML algorithms were compared with each other using models that included all radiological features and the top 14 radiomics parameters, GB achieved the highest specificity, reaching 99%. The best overall model performance in terms of accuracy was observed with LR and SVM models. Notably, all ML algorithms demonstrated AUC values exceeding 0.91, outperforming radiological assessment alone.

### Performance comparison on the 7<sup>th</sup> post-contrast images

On multivariate LR, radiological features alone again demonstrated limited specificity, as the same radiological parameters were used across phases. The addition of radiomics features improved model performance, with combined radiomics models increasing specificity from 78% to 93% and accuracy from 82% to 91%.

ML models trained on radiological features alone achieved high specificity (up to 96%) and an AUC of 0.92, consistent with early-phase results. The highest overall performance was obtained with the combined model integrating radiological features and 14 radiomics features selected using RFE (AUC, 0.93). Among ML algorithms, GB achieved the highest specificity, whereas LR and SVM demonstrated the best overall accuracy.

Model performances are summarized in Tables 3-5.

## Discussion

This study aimed to improve the diagnostic performance of breast MRI in the evaluation of NME by integrating volumetric 3D radiomics features and using ML algorithms with both early (2<sup>nd</sup>) and late (7<sup>th</sup>) post-contrast series. We found that the integration of radiomics and ML substantially improved diagnostic performance compared with radiological assessment alone for the evaluation of NME. Notably, consistent improvements in accuracy and specificity were observed across both early and late phases, with com-

pared radiomics-ML models demonstrating the highest overall performance. These findings support the clinical value of radiomics and ML by improving specificity and accuracy. In the clinical workflow, this can enable more confident decision-making in NME assessment.

NME remains one of the most challenging entities in breast MRI, representing a major source of false-positive findings despite high sensitivity. This challenge is largely due to substantial benign-malignant overlap and significant intra- and inter-observer vari-

ability in interpretation.<sup>1-6</sup> Radiomics offers a comprehensive approach to tissue characterization beyond visual interpretation. By providing objective and standardized measurements, radiomics can mitigate the impact of reader experience and reduce inter- and intra-observer variability, enabling a more reproducible assessment. Furthermore, ML algorithms have the potential to improve diagnostic accuracy and reduce subjectivity.

One of the main contributions of our study was the 3D volumetric evaluation. Although NMEs do not exhibit classic mass

**Table 1.** Comparison of distribution, internal enhancement, laterality, background parenchymal enhancement, breast density, T2-weighted signal, patient age, and maximum lesion diameter of non-mass enhancement between the benign and malignant groups

Radiological features	Groups	Benign	Malign	P
Distribution	Focal	36 (41.9)	0 (0.0)	0.000
	Linear-straight	17 (19.8)	0 (0.0)	
	Linear-branching	12 (14.0)	2 (8.3)	
	Segmental	16 (18.6)	13 (54.2)	
	Regional	3 (3.5)	7 (29.2)	
	Multiple regional	1 (1.2)	0 (0.0)	
Internal enhancement	Diffuse	1 (1.2)	2 (8.3)	0.000
	Homogeneous	32 (37.2)	1 (4.2)	
	Heterogeneous	10 (11.6)	12 (50.0)	
	Clumped	38 (44.2)	5 (20.8)	
Laterality	Clustered ring	6 (7.0)	6 (25.0)	0.021
	L	48 (55.8)	7 (29.2)	
	R	38 (44.2)	17 (70.8)	
Background parenchymal enhancement	Minimal	7 (8.1)	2 (8.3)	0.962
	Mild	21 (24.4)	5 (20.8)	
	Moderate	31 (36.0)	10 (41.7)	
	Marked	27 (31.4)	7 (29.2)	
Breast density	A	1 (1.2)	1 (4.2)	0.772
	B	18 (20.9)	4 (16.7)	
	C	52 (60.5)	15 (62.5)	
	D	15 (17.4)	4 (16.7)	
T2W signal	Hypointense	33 (38.4)	10 (41.7)	0.228
	Isointense	36 (41.9)	6 (25.0)	
	Hyperintense	17 (19.8)	8 (33.3)	
Age		47.87 ± 9.69	49.48 ± 11.58	0.699
Max diameter (mm)		17.24 ± 11.73	45.17 ± 30.5	0.000

**Table 2.** Number of statistically significant radiomics features ( $P < 0.05$ ) identified out of 123 features on the 2<sup>nd</sup> and 7<sup>th</sup> post-contrast images

Feature group	2 <sup>nd</sup> post-contrast	7 <sup>th</sup> post-contrast
Morphological	14/14	14/14
Intensity-based	30/54	23/54
Second-order texture	48/55	43/55
Total	92/123	80/123

characteristics,<sup>1</sup> they occupy a 3D space within the breast parenchyma, and their full spatial extent may not be adequately captured by 2D analysis. We hypothesize that 2D evaluation does not maximize diagnostic performance in NMEs.

Previous studies have demonstrated the potential value of radiomics in NME assessment in this context. Li et al.<sup>19,20</sup> reported that radiomics models integrating clinical and MRI features achieved high sensitivity and specificity and improved the characterization of BI-RADS 4 NMEs without the need for additional DWI. Zhou et al.<sup>21</sup> compared radiomics and deep learning approaches and showed that an SVM classifier achieved accuracies of 80.4% in the training set and 77.5% in the test set. Kayadibi et al.<sup>22</sup> similarly demonstrated that a clinical–radiomics ML model outperformed radiologist assessment alone in differentiating malignant lesions from idiopathic granulomatous mastitis. Unlike previous studies that predominantly relied on 2D slice-based analysis, the present study employed full volumetric segmentation of NME. Our 3D radiomics framework enables a more comprehensive characterization of lesion architecture and heterogeneity.

The rationale for analyzing both early (2<sup>nd</sup>) and late (7<sup>th</sup>) post-contrast phases in this study is based on the limited diagnostic value of kinetic assessment alone in NMEs. Unlike mass lesions, kinetic curve analysis in NMEs shows substantial overlap between benign and malignant entities, with relatively low positive predictive values reported in the literature.<sup>24</sup> Previous studies have demonstrated that persistent or plateau-type enhancement patterns are common in both benign and malignant NMEs, whereas wash-out kinetics are relatively uncommon and unreliable, particularly in DCIS and invasive lobular carcinoma.<sup>9,10,25</sup> Moreover, kinetic parameters that are useful for differentiating mass lesions are ineffective in reliably discriminating NMEs.<sup>26</sup> Based on this limitation, we aimed to capture the full biological behavior of NMEs by performing radiomics analysis on both early and late post-contrast images. This approach also allowed us to compare the early and late post-contrast phases. Notably, a greater number of radiomics features were significantly associated with benign–malignant differentiation on the early post-contrast images compared with the late phase. Consistently, ML models incorporating radiomics features demonstrated higher diagnostic accuracy on the 2<sup>nd</sup> post-contrast series compared with the 7<sup>th</sup> series (Tables 3–5). Our findings support the use of early

**Table 3.** Performance comparison of logistic regression and machine learning models using radiological and radiomics features on the 2<sup>nd</sup> post-contrast images

	Model	Sensitivity	Specificity	Accuracy	AUC
<b>Logistic regression</b>	Radiological parameters	96%	78%	82%	0.90
	Radiomics general	83%	98%	93%	0.96
	Radiomics–morphological	96%	79%	86%	0.94
	Radiomics–histogram	92%	84%	88%	0.93
	Radiomics–second order statistics	88%	95%	84%	0.96
<b>Machine learning</b>	Radiological parameters	57%	96%	88%	0.92
	Radiomics–morphological	71%	95%	90%	0.92
	Radiomics–histogram	78%	80%	80%	0.88
	Radiomics–second order statistics	91%	79%	81%	0.93
	14 radiomics parameters (RFE)	75%	95%	91%	0.95
	Radiological parameters + best 14 radiomics	83%	90%	88%	0.95

AUC, area under the curve; RFE, recursive feature elimination.

**Table 4.** Performance comparison of logistic regression and machine learning models using radiological and radiomics features on the 7<sup>th</sup> post-contrast images

	Model	Sensitivity	Specificity	Accuracy	AUC
<b>Logistic regression</b>	Radiological parameters	96%	78%	82%	0.90
	Radiomics general	92%	93%	91%	0.95
	Radiomics–morphological	95%	75%	85%	0.92
	Radiomics–histogram	92%	82%	85%	0.92
	Radiomics–second order statistics	54%	58%	22%	0.50
<b>Machine learning</b>	Radiological parameters	57%	96%	88%	0.92
	Radiomics–morphological	66%	94%	88%	0.89
	Radiomics–histogram	58%	86%	80%	0.87
	Radiomics–second order statistics	92%	76%	79%	0.91
	14 radiomics parameters (RFE)	67%	87%	83%	0.91
	Radiological parameters + best 14 radiomics	88%	86%	87%	0.93

AUC, area under the curve; RFE, recursive feature elimination.

**Table 5.** Performance comparison of machine learning algorithms on the 2<sup>nd</sup> and 7<sup>th</sup> post-contrast images

	Model	Sensitivity	Specificity	Accuracy	AUC
<b>2<sup>nd</sup> post-contrast</b>	Logistic regression	83%	90%	88%	0.95
	Random forest	67%	95%	89%	0.93
	Support vector machine	75%	95%	91%	0.95
	Gradient boost	54%	99%	88%	0.91
<b>7<sup>th</sup> post-contrast</b>	Logistic regression	88%	86%	87%	0.93
	Random forest	53%	87%	80%	0.88
	Support vector machine	54%	91%	83%	0.93
	Gradient boost	46%	95%	85%	0.90

AUC, area under the curve.

post-contrast imaging as a more informative phase for radiomics-based NME analysis in improving diagnostic performance.

In the early and late phases, 92 of 123 and 80 of 123 radiomics features, respectively, were statistically significant. These findings suggest that radiomics may provide an extensive set of quantitative data that could complement conventional radiological and clinical parameters; however, these results should be interpreted with caution, as no formal correction for multiple comparisons was applied. Using the obtained data, diagnostic models were constructed by combining radiological, clinical, and radiomics data.

Our results confirm that radiomics improves diagnostic model performance beyond conventional radiological assessment. Combining morphological, intensity-based, and second-order texture features provided better performance than using individual feature groups alone, indicating that comprehensive volumetric characterization is more beneficial. Importantly, ML applied even to radiological features alone substantially increased specificity, highlighting its ability to reduce false-positive interpretations. Among the evaluated algorithms, GB consistently showed the highest specificity, whereas LR and SVM maintained strong overall accuracy.

As a result of the phase-specific analyses we performed, the model performances obtained on the early and late phases were comparable, as demonstrated in the Results section. This consistency suggests that the use of radiomics and ML algorithms in NME assessment is reproducible and can be reliably integrated into the clinical workflow. Future studies should focus on accelerating the integration of radiomics into routine clinical workflows with automatic segmentation to facilitate broader and more practical clinical adoption.

This study has several limitations. First, its retrospective design may have introduced selection bias. Second, MRI-guided biopsy was not available at our institution; therefore, a substantial proportion of benign lesions were classified based on imaging follow-up rather than histopathological confirmation, resulting in a heterogeneous reference standard. Third, the relatively small number of malignant lesions, together with class imbalance between benign and malignant cases, may have affected model stability and inflated some performance metrics. Fourth, although internal evaluation was

performed using 5-fold cross-validation, RFE was applied before cross-validation, which may have increased the risk of information leakage and overfitting; therefore, the findings should be considered exploratory and preliminary. Fifth, calibration analysis was not performed, so the reliability of predicted probabilities could not be assessed. Finally, segmentation was performed manually, which may limit reproducibility and practical applicability despite quality control by a second radiologist in a subset of cases. Further studies with larger cohorts, external validation, calibration analysis, and automated or semi-automated segmentation are needed.

In conclusion, volumetric assessment of NME using radiomics and ML showed promising results for improving diagnostic accuracy and specificity. These findings suggest a potential role for radiomics–ML models in increasing diagnostic confidence and reproducibility and possibly in reducing unnecessary biopsies; however, these implications should be considered preliminary. From a clinical perspective, the proposed radiomics–ML model should not be regarded as a replacement for radiologist interpretation, but rather as a decision-support tool for the assessment of NME on breast MRI. A practical workflow may involve lesion detection and conventional radiological assessment first, followed by volumetric segmentation and model-based risk estimation in indeterminate cases. Before routine implementation, further steps are required, including external validation in larger multicenter cohorts, standardization of imaging protocols, calibration assessment, and the development of automated or semi-automated segmentation tools to improve feasibility in daily practice.

#### Footnotes

#### Conflict of interest disclosure

Işıl Başara Akin, MD, serves as Section Editor for Diagnostic and Interventional Radiology. She had no involvement in the peer review of this article and had no access to information regarding its peer review. The other authors declare no conflicts of interest.

**Supplementary:** <https://d2v96fxpocvxx.cloudfront.net/beb8919b-f013-4ea1-b1c8-40332e840fe1/content-images/730bc-1ca-e8a9-4643-9f19-6d57f16d2bf0.pdf>

**Appendix:** <https://d2v96fxpocvxx.cloudfront.net/beb8919b-f013-4ea1-b1c8-40332e840fe1/content-images/97689d0e-5b7a-4103-8a2a-4de870591ee8.pdf>

## References

1. Seely JM, Bissell MB. BI-RADS v2025: a welcome update. *AJR Am J Roentgenol*. 2026. [Crossref]
2. Berg WA, Gutierrez L, NessAiver MS, et al. Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. *Radiology*. 2004;233(3):830-849. [Crossref]
3. Baltzer PA, Benndorf M, Dietzel M, Gajda M, Runnebaum IB, Kaiser WA. False-positive findings at contrast-enhanced breast MRI: a BI-RADS descriptor study. *AJR Am J Roentgenol*. 2010;194(6):1658-1663. [Crossref]
4. Lunkiewicz M, Forte S, Freiwald B, Singer G, Leo C, Kubik-Huch RA. Interobserver variability and likelihood of malignancy for fifth edition BI-RADS MRI descriptors in non-mass breast lesions. *Eur Radiol*. 2020;30(1):77-86. [Crossref]
5. Torous VF, Resteghini NA, Phillips J, et al. Histopathologic correlates of nonmass enhancement detected by breast magnetic resonance imaging. *Arch Pathol Lab Med*. 2021;145(10):1264-1269. [Crossref]
6. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol*. 2005;23(33):8469-8476. [Crossref]
7. Shao Z, Wang H, Li X, Liu P, Zhang S, Cao S. Morphological distribution and internal enhancement architecture of contrast-enhanced magnetic resonance imaging in the diagnosis of non-mass-like breast lesions: a meta-analysis. *Breast J*. 2013;19(3):259-268. [Crossref]
8. Peters NH, Borel Rinkes IH, Zuihthoff NP, Mali WP, Moons KG, Peeters PH. Meta-analysis of MR imaging in the diagnosis of breast lesions. *Radiology*. 2008;246(1):116-124. [Crossref]
9. Yang QX, Ji X, Feng LL, et al. Significant MRI indicators of malignancy for breast non-mass enhancement. *J Xray Sci Technol*. 2017;25(6):1033-1044. [Crossref]
10. Aydin H. The MRI characteristics of non-mass enhancement lesions of the breast: associations with malignancy. *Br J Radiol*. 2019;92(1096):20180464. [Crossref]
11. Yuen S, Uematsu T, Masako K, Uchida Y, Nishimura T. Segmental enhancement on breast MR images: differential diagnosis and diagnostic strategy. *Eur Radiol*. 2008;18(10):2067-2075. [Crossref]
12. Tozaki M, Igarashi T, Fukuda K. Breast MRI using the VIBE sequence: clustered ring enhancement in the differential diagnosis of lesions showing non-masslike enhancement. *AJR Am J Roentgenol*. 2006;187(2):313-321. [Crossref]

13. Ballesio L, Di Pastena F, Gigli S, et al. Non mass-like enhancement categories detected by breast MRI and histological findings. *Eur Rev Med Pharmacol Sci*. 2014;18(6):910-917. [\[Crossref\]](#)
14. Asada T, Yamada T, Kanemaki Y, Fujiwara K, Okamoto S, Nakajima Y. Grading system to categorize breast MRI using BI-RADS 5th edition: a statistical study of non-mass enhancement descriptors in terms of probability of malignancy. *Jpn J Radiol*. 2018;36(3):200-208. [\[Crossref\]](#)
15. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. *Radiology*. 2021;298(3):505-516. Erratum in: *Radiology*. 2021;299(2):E256. [\[Crossref\]](#)
16. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to radiomics. *J Nucl Med*. 2020;61(4):488-495. [\[Crossref\]](#)
17. Rogers W, Thulasi Seetha S, Refaee TAG, et al. Radiomics: from qualitative to quantitative imaging. *Br J Radiol*. 2020;93(1108):20190948. [\[Crossref\]](#)
18. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Radiographics*. 2017;37(2):505-515. [\[Crossref\]](#)
19. Li Y, Yang ZL, Lv WZ, et al. Non-mass enhancements on DCE-MRI: development and validation of a radiomics-based signature for breast cancer diagnoses. *Front Oncol*. 2021;11:738330. [\[Crossref\]](#)
20. Li Y, Yang Z, Lv W, et al. Role of combined clinical-radiomics model based on contrast-enhanced MRI in predicting the malignancy of breast non-mass enhancements without an additional diffusion-weighted imaging sequence. *Quant Imaging Med Surg*. 2023;13(9):5974-5985. [\[Crossref\]](#)
21. Zhou J, Liu YL, Zhang Y, et al. BI-RADS reading of non-mass lesions on DCE-MRI and differential diagnosis performed by radiomics and deep learning. *Front Oncol*. 2021;11:728224. [\[Crossref\]](#)
22. Kayadibi Y, Saracoglu MS, Kurt SA, et al. Differentiation of malignancy and idiopathic granulomatous mastitis presenting as non-mass lesions on MRI: radiological, clinical, radiomics, and clinical-radiomics models. *Acad Radiol*. 2024;31(9):3511-3523. [\[Crossref\]](#)
23. Kocak B, Akinci D'Antonoli T, Mercaldo N, et al. METHodological RadiomIcs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMI. *Insights Imaging*. 2024;15(1):8. [\[Crossref\]](#)
24. Balci P. Non-mass findings on MRI and differential diagnosis. *Turk Radiol Semin*. 2023;11(3):312-325. [\[Crossref\]](#)
25. Kuhl C. The current status of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology*. 2007;244(2):356-378. [\[Crossref\]](#)
26. Jansen SA, Fan X, Karczmar GS, et al. DCEMRI of breast lesions: is kinetic analysis equally effective for both mass and nonmass-like enhancement? *Med Phys*. 2008;35(7):3102-3109. [\[Crossref\]](#)