



# Multiparametric MRI-based radiomics nomogram for preoperative prediction of lymphovascular invasion and clinical outcomes in patients with breast invasive ductal carcinoma

Junjie Zhang<sup>1</sup> · Guanghui Wang<sup>2</sup> · Jialiang Ren<sup>3</sup> · Zhao Yang<sup>1</sup> · Dandan Li<sup>1</sup> · Yanfen Cui<sup>1</sup> · Xiaotang Yang<sup>1</sup>

Received: 20 July 2021 / Revised: 1 December 2021 / Accepted: 4 December 2021 / Published online: 20 January 2022  
© The Author(s), under exclusive licence to European Society of Radiology 2022

## Abstract

**Objective** To develop a multiparametric MRI-based radiomics nomogram for predicting lymphovascular invasion (LVI) status and clinical outcomes in patients with breast invasive ductal carcinoma (IDC).

**Methods** A total of 160 patients with pathologically confirmed breast IDC (training cohort:  $n = 112$ ; validation cohort:  $n = 48$ ) who underwent preoperative breast MRI were included. Imaging features were extracted from T2-weighted imaging (T2WI), apparent diffusion coefficient (ADC) maps, and contrast-enhanced T1-weighted imaging (cT1WI) sequences. A four-step procedure was applied for feature selection and radiomics signature building. Univariate and multivariate logistic regression analyses were conducted to identify the features associated with LVI, which were then incorporated into the radiomics nomogram. The performance of the nomogram was evaluated by its discrimination, calibration, and clinical usefulness. Kaplan–Meier survival curves based on the two radiomics models were used to estimate disease-free survival (DFS).

**Results** The fusion radiomics signature of the T2WI, cT1WI, and ADC maps achieved a better predictive efficacy for LVI than either of them alone. The proposed radiomics nomogram, incorporating the fusion radiomics signature and MRI-reported peritumoral edema, showed satisfactory capabilities of calibration and discrimination in both training and validation datasets, with AUCs of 0.919 (95% CI: 0.871–0.967) and 0.863 (95% CI: 0.726–0.999), respectively. The radiomics signature and nomogram-defined high-risk groups had a shorter DFS than those in the low-risk groups (both  $p < 0.05$ ). Higher Rad-scores were independently associated with a worse DFS in the whole cohort ( $p < 0.05$ ).

**Conclusions** The proposed nomogram, incorporating multiparametric MRI-based radiomics signature and MRI-reported peritumoral edema, achieved a satisfactory preoperative prediction of LVI and clinical outcomes in IDC patients.

## Key Points

- The fusion radiomics signature of the T2WI, cT1WI, and ADC maps achieved a better predictive efficacy for LVI than either of them alone.
- The proposed nomogram achieved a favorable prediction of LVI in IDC patients with AUCs of 0.919 and 0.863 in the training and validation datasets, respectively.
- The radiomics model could classify patients into high- and low-risk groups with significant differences in DFS.

**Keywords** Magnetic resonance imaging · Nomograms · Lymphovascular invasion · Breast neoplasms

Junjie Zhang and Guanghui Wang contributed equally to this work.

✉ Yanfen Cui  
yanfen210@126.com

✉ Xiaotang Yang  
yangxt210@126.com

<sup>1</sup> Department of Radiology, Shanxi Province Cancer Hospital, Shanxi Medical University, Taiyuan 030013, China

<sup>2</sup> Department of Breast Surgery, Guizhou Provincial People's Hospital, Guiyang 550002, China

<sup>3</sup> GE Healthcare China, Beijing, China

## Abbreviations

AIC	Akaike information criterion
AVS	Adjacent vessel sign
BCS	Breast-conserving surgery
cT1WI	Contrast-enhanced T1-weighted imaging
DCA	Decision curve analysis
DFS	Disease-free survival
ICC	Interclass correlation coefficient
IDC	Invasive ductal carcinoma
LASSO	Least absolute shrinkage and selection operator
LVI	Lymphovascular invasion

mrALN	MRI-reported axillary lymph nodes
NME	Nonmass enhancement
T2WI	T2-weighted imaging
TIC	Time-intensity curve

## Introduction

Breast cancer has become the most commonly diagnosed cancer, as well as the leading cause of cancer-related deaths in women worldwide [1]. In these patients, it is the metastases at distant sites, rather than the primary tumor, that are the main cause of death [2]. One of the key factors of metastases is lymphovascular invasion (LVI), which was defined as the infiltration of neoplastic cells within blood vessels or lymphatic lumens [3]. Notably, it is one of the critical steps in the metastasis, and it has been proved to be an independent prognostic factor for local recurrence (LR) or distant metastasis (DM) of breast cancer [4]. At present, LVI can only be diagnosed with a surgical specimen postoperatively [5]. Preoperative knowledge of LVI in breast cancer may assist clinicians in making more reliable clinical decisions, including type of surgery, surgical margin determination, and guiding more aggressive neoadjuvant treatment protocols. However, preoperative prediction of LVI is still a main challenge [6].

Clinicopathologic risk indicators related to LVI of breast cancer have been studied, but the results are inconsistent [7, 8]. MRI is an important diagnostic tool for tumor detection and prediction of neoadjuvant therapy or prognosis in patients with breast cancer, owing to the advantages of high-resolution and multiparameter scanning [9]. Previous studies have shown that some MRI features, such as peritumoral edema, adjacent vessel sign (AVS), and tumor apparent diffusion coefficient (ADC) values were associated with LVI [10–14]. However, these features have not been extensively utilized for prediction of LVI, and they should be further assessed clinically.

Radiomics can capture underlying pathophysiological information, and it has noticeably attracted scholars' attention [15, 16]. The association between intratumor heterogeneity and LVI has been previously reported [17–20]. To date, few studies have evaluated the application of radiomics to predict LVI in patients with breast cancer [21], and no study has specifically concentrated on the potential value of multiparametric MRI-based radiomics to reliably predict LVI and outcomes.

Therefore, the present study aimed to investigate whether the radiomics nomogram, integrating clinical factors, MRI features, and radiomics features extracted from multiparametric MRI could be used to predict LVI and outcomes in patients with invasive breast cancer.

## Materials and methods

### Patients

This retrospective study was approved by the institutional ethics review board of our hospital and the requirement of informed consent was waived. A total of 538 consecutive female patients with breast invasive ductal carcinoma (IDC), who underwent mastectomy or breast-conserving surgery from March 2014 to December 2017, were recruited. Patients were excluded if one of the following criteria was met: LVI status was not explicitly reported pathologically, patients with a history of other malignant tumors, undergoing neoadjuvant chemotherapy preoperatively, undergoing vacuum-assisted excisional biopsy before the MRI examination, poor quality of imaging, or unavailable follow-up data. Finally, a total of 160 patients with IDC were included and were randomly assigned into training dataset ( $n = 112$ ) and validation dataset ( $n = 48$ ) at a ratio of 7:3.

### Histopathology

All patients underwent surgical treatment within 2 weeks after MRI examination. The pathological type, grade, and LVI status were recorded. LVI was defined as positive when there were cancer cells in lymphatic vessels or small blood vessels outside the tumor. Immunohistochemistry (IHC) was performed to detect the expression levels of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER-2), and Ki-67 expression status. The cutoff value for ER and PR was set at 1%, and the cutoff value for Ki-67 was 20%. Concerning the HER-2 status, according to the American Society of Clinical Oncology (ASCO)/College of American Pathologists (CAP) guidelines [22], tumors with an IHC staining score of 0 or 1+ were defined as HER-2-negative and those with an IHC staining score of 3+ as HER-2-positive. For tumors with an IHC staining score of 2+, fluorescence in situ hybridization (FISH) was further performed to confirm the HER-2 status. A non-amplified FISH result denotes the HER-2 status as negative, and an amplified result denotes the HER-2 status as positive. Breast cancer can be classified into four main immunohistochemical subtypes, namely luminal A, luminal B, HER2-enriched, and triple negative subtypes according to the 2013 St. Gallen Consensus Conference [23].

### MRI protocol and radiologic evaluation

All MRI examinations were performed on Achieva 3.0-T system (Philips Healthcare) with an eight-channel phased array breast coil. The routine protocols were composed

of turbo spin-echo T1-weighted imaging (T1WI), fat-suppressed spin-echo T2-weighted imaging (T2WI), echo-planar diffusion-weighted imaging (DWI), and dynamic contrast-enhanced MRI (DCE-MRI). Imaging acquisition parameters were detailed in Appendix E1. For patients with multiple lesions, the largest lesion was selected for radiomics analysis.

MR images were reviewed independently by two breast radiologists with 15 and 10 years of MR experience respectively, and consensus was reached by discussion in event of disagreement. Both radiologists were blinded to all histopathological data. Those radiologists assessed the following imaging features of tumors: tumor size, lesion type, tumor margin, time-intensity curve (TIC) pattern, peritumoral edema, AVS, and MRI-reported axillary lymph node (mrALN) status. Peritumoral edema positive was defined as the higher signal intensity surrounding the tumor compared with that of surrounding breast tissue or almost equal to that of water by visual evaluation on T2W images [24]. The presence of the AVS was defined as the presence of vessel either entering the lesion or contacting the edge of the lesion on the contrast T1W images [10]. Positive-MRI findings of ALN were accompanied with a short-axis diameter exceeding 10 mm, a ratio of the longest axis to the shortest axis was  $< 1.5$ , loss of fatty hilum, and eccentric cortical thickening [25].

### MRI segmentation and radiomics features extraction

Regions of interests (ROIs) were manually drawn on the transverse T2WI, DWI, and cT1WI (the early enhanced phase at 60 s after contrast injection), along the edge of the lesion slice-by-slice for each patient, by the above two radiologists using ITK-SNAP 3.8 software (<http://www.itksnap.org>). The ROIs on DWI were automatically mapped to ADC map.

Image preprocessing and feature extraction were performed using open-source Pyradiomics software (<http://PyRadiomics.readthedocs.io/en/latest/>). Details of the feature extraction algorithms are presented in Appendix E2.

### Radiomics feature selection and radiomics signature building

We devised a four-step procedure for selecting robust radiomics features. Firstly, the intraclass correlation coefficient (ICC) was calculated to explore the stability of radiomics features, and features with an ICC  $< 0.80$  were excluded. Secondly, the univariate analysis was performed to select LVI-related features (with  $p < 0.01$ ). Thirdly, the Spearman's correlation analysis with  $r \geq 0.90$  was used to eliminate the redundancy. Finally, we used the least absolute shrinkage and selection operator (LASSO) method to select the most

significant features in the training dataset. Multivariate logistic regression analysis was then conducted to develop a radiomics signature using the Akaike information criterion (AIC) as the stopping rule for predicting LVI.

### Development, performance, and validation of the radiomics nomogram

Univariate analysis was used to assess the association between the clinico-radiological (clinicopathological and radiological) characteristics and LVI status, and significant risk factors were considered for the multivariate logistic regression analysis to develop the clinico-radiological model. Similarly, the radiomics signature and the abovementioned clinico-radiological signatures were then tested using univariate and multivariable logistic regression analyses to construct the radiomics nomogram for prediction of LVI in the training dataset.

Receiver operating characteristic (ROC) curves and the area under the curve (AUC) values were used to assess the discrimination performance of the established models for LVI prediction. Calibration curve was adopted to investigate the calibration of the radiomics nomogram in the training and validation datasets. The Hosmer–Lemeshow test was used to assess the goodness-of-fit of the radiomics nomogram. Decision curve analysis (DCA) was conducted to estimate the clinical usefulness of the radiomics nomogram by calculating the net benefits at different values of threshold probability.

### Statistical analysis

The statistical analysis was conducted using the SPSS 22.0 software (IBM Corp.) and R software version 3.6.3 (<http://www.R-project.org>, Appendix E3). Clinico-radiological features of patients in LVI-positive and LVI-negative groups and in the training and validation datasets were compared by the Student *t*-test or the Mann–Whitney *U* test for continuous variables and Chi-squared or the Fisher exact test for categorical variables. Disease-free survival (DFS) was the outcome of interest, which was defined as the time from the date of surgery to that of the first LR or DM or death from any cause. The association of the radiomics nomogram with DFS was analyzed by using the Kaplan–Meier method and compared by the log-rank test. The optimum cutoff values to classify patients into high-risk and low-risk groups, according to the radiomics signature or nomogram, were identified using the maximally selected rank statistical method. Univariate and multivariate logistic regression analyses with the Cox proportional-hazards regression determined the predictors of DFS. Variables that reached a statistical significance in the univariate analysis were considered for the multivariate logistic regression analysis. A two-sided  $p < 0.05$  was considered statistically significant.

## Results

### Clinico-radiological characteristics

No significant difference was found in all the clinico-radiological features between the training and validation datasets ( $p=0.056$ – $0.988$ ) (Table S1). The clinico-radiological features of patients in LVI-positive and LVI-negative groups in the training and validation datasets are shown in Table 1. The rates of LVI were 36.7% (41 of 112) and 31.3% (15 of 48) in the training and validation datasets, respectively.

The univariate analysis showed that tumor size, lesion type, tumor margin, peritumoral edema, AVS, and mrALN status were significantly correlated to LVI ( $p < 0.05$ ) (Table S2). The multivariate logistic regression analysis revealed that peritumoral edema (odds ratio [OR] = 3.500; 95% confidence interval [CI]: 1.611–7.604;  $p = 0.002$ ), AVS (OR = 2.266; 95%CI: 1.025–5.010;  $p = 0.043$ ), and mrALN status (OR = 3.503; 95%CI: 1.625–7.552;  $p = 0.001$ ) remained independent predictors of LVI (Table 2). A clinico-radiological model was then established based on the abovementioned independent predictors of LVI. Representative MR images of LVI-positive and LVI-negative of IDC are presented in Figure S1.

### Radiomics feature selection, radiomics signature building, and validation

In this study, a total of 3111 imaging features for each patient were extracted, 1037 features were from each of the ADC, cT1WI, and T2WI sequences. We selected 930, 897, and 976 features with high stability and reproducibility ( $ICC > 0.80$ ) from these features for the above sequences, respectively. Afterward, a total of 1228 features were selected by the univariate analysis. After eliminating redundancy by applying the Pearson correlation analysis, 96 features remained, and were then subjected to the LASSO method. Finally, 3 ADC features, 2 cT1WI features, and 7 T2WI features were selected. The Rad-score calculation formulas of each sequence were constructed with the corresponding selected features, respectively (Appendix E4).

For the fusion radiomics signature derived from the abovementioned sequences, eight radiomics features were selected after the multivariate logistic regression analysis, including two wavelet features and six Laplacian of Gaussian (LoG) features. All these eight features were significantly different between the LVI-positive and LVI-negative groups (all  $p < 0.05$ ; Fig. 1). The Rad-score calculation formula of the fusion radiomics signature model was constructed (Appendix E4). The distributions of the Rad-scores for each patient in the training and validation datasets are shown in Fig. 2a and b. The Rad-scores in the LVI-positive group were significantly higher than those in

the LVI-negative group in the training and validation datasets (Fig. 2c and d).

The prediction performances of the single radiomics signature and the fusion radiomics signature for predicting LVI status are presented in Fig. 3 and Table 3. The radiomics signature derived from ADC, cT1WI, and T2WI sequences yielded AUC values of 0.804 (95%CI: 0.719–0.889), 0.800 (95%CI: 0.717–0.883), and 0.860 (95%CI: 0.790–0.930) in the training dataset, which were then confirmed in the validation dataset with the AUC values of 0.762 (95%CI: 0.600–0.924), 0.713 (95%CI: 0.558–0.868), and 0.844 (95%CI: 0.705–0.984), respectively. The fusion radiomics signature yielded the highest AUC values of 0.905 (95%CI: 0.850–0.960) and 0.848 (95%CI: 0.712–0.985) in the training and validation datasets, respectively.

### Development, performance, and validation of the radiomics nomogram

On multivariate logistic regression analysis, the radiomics signature and peritumoral edema were identified as independent risk factors for predicting LVI status. The radiomics nomogram incorporating the abovementioned independent predictors showed the highest discrimination ability for predicting LVI, with AUC values of 0.919 (95%CI: 0.871–0.967) and 0.863 (95%CI: 0.726–0.999) in the training and validation datasets, respectively (Figs. 3 and 4, Figure S2, and Table 3).

The calibration curve of the radiomics nomogram showed a good agreement between the predicted LVI probabilities and the actual LVI status in both training and validation datasets (Fig. 5a and b). The DCA showed that the radiomics nomogram had higher net benefit than the radiomics signature and the clinico-radiological model in predicting LVI in the two datasets (Fig. 5c and d).

### Predictors of survival

The median follow-up period was 41 months (range, 5–75 months). LR, DM, or deaths from any cause occurred in 25 patients (15.6%) at a median follow-up period of 16 months (range, 5–36 months).

We identified the optimal radiomics signature or nomogram-defined score for prediction of DFS as 0.711 and 0.581, respectively. Accordingly, all patients were divided into high-risk and low-risk groups, and the Kaplan–Meier survival curves showed that higher radiomics signature or nomogram-defined scores were significantly associated with a worse DFS in the entire cohort (hazard ratio [HR], 2.950; 95%CI: 1.339–6.501,  $p = 0.005$  versus HR, 3.160; 95%CI: 1.434–6.963,  $p = 0.003$ ) (Fig. 6). Table S3 summarizes the results of the univariate Cox regression

**Table 1** Comparisons of clinico-radiological features between patients with LVI and without LVI in the training and validation datasets

Characteristics	Training dataset (n = 112)		p value	Validation dataset (n = 48)		p value
	LVI present (n = 41)	LVI absent (n = 71)		LVI present (n = 15)	LVI absent (n = 33)	
Age, years	49.5 ± 10.3	48.0 ± 10.2	0.455	49.6 ± 8.0	48.5 ± 8.8	0.684
Menopausal status			0.389			0.327
Premenopausal	22 (53.7%)	44 (62.0%)		10 (66.7%)	17 (51.5%)	
Postmenopausal	19 (46.3%)	27 (38.0%)		5 (33.3%)	16 (48.5%)	
Tumor size <sup>a</sup> , cm	2.8 (2.0–4.4)	1.8 (1.5–2.5)	0.000 <sup>b</sup>	2.9 (1.6–4.0)	2.3 (1.6–2.7)	0.095 <sup>b</sup>
Lesion type			0.008			0.084 <sup>c</sup>
Mass	29 (70.7%)	64 (90.1%)		12 (80.0%)	32 (97.0%)	
NME	12 (29.3%)	7 (9.9%)		3 (20.0%)	1 (3.0%)	
Margin			0.013			0.409 <sup>c</sup>
Circumscribed	3 (7.3%)	19 (26.8%)		1 (6.7%)	6 (18.2%)	
Not circumscribed	38 (92.7%)	52 (73.2%)		14 (93.3%)	27 (81.8%)	
TIC pattern			0.927 <sup>c</sup>			0.537 <sup>c</sup>
Type I	1 (2.4%)	2 (2.8%)		0 (0.0%)	1 (3.0%)	
Type II	25 (61.0%)	45 (63.4%)		8 (53.3%)	19 (57.6%)	
Type III	15 (36.6%)	24 (33.8%)		7 (46.7%)	13 (39.4%)	
Peritumoral edema			0.000			0.031
Negative	11 (26.8%)	50 (70.4%)		5 (33.3%)	22 (66.7%)	
Positive	30 (73.2%)	21 (29.6%)		10 (66.7%)	11 (33.3%)	
AVS			0.000			0.283 <sup>c</sup>
Negative	17 (41.5%)	55 (77.5%)		10 (66.7%)	27 (81.8%)	
Positive	24 (58.5%)	16 (22.5%)		5 (33.3%)	6 (18.2%)	
mrALN status			0.000			0.016
Negative	16 (39.0%)	56 (78.9%)		6 (40.0%)	25 (75.8%)	
Positive	25 (61.0%)	15 (21.1%)		9 (60.0%)	8 (24.2%)	
ER status			0.766			1.000 <sup>c</sup>
Negative	15 (36.6%)	24 (33.8%)		3 (20.0%)	7 (21.2%)	
Positive	26 (63.4%)	47 (66.2%)		12 (80.0%)	26 (78.8%)	
PR status			0.440			0.809
Negative	21 (51.2%)	31 (43.7%)		6 (40.0%)	12 (36.4%)	
Positive	20 (48.8%)	40 (56.3%)		9 (60.0%)	21 (63.6%)	
HER2 status			0.685			0.283 <sup>c</sup>
Negative	32 (78.0%)	53 (74.6%)		10 (66.7%)	27 (81.8%)	
Positive	9 (22.0%)	18 (25.4%)		5 (33.3%)	6 (18.2%)	
Ki-67 index(%)			0.224			0.727 <sup>c</sup>
< 20	4 (9.8%)	13 (18.3%)		3 (20.0%)	10 (30.3%)	
≥ 20	37 (90.2%)	58 (81.7%)		12 (80.0%)	23 (69.7%)	
Molecular subtype			0.542			0.983 <sup>c</sup>
Luminal A	3 (7.3)	12 (16.9)		3 (20.0)	10 (30.3)	
Luminal B	23 (56.1)	37 (52.1)		9 (60.0)	16 (48.5)	
HER2 positive	8 (19.5)	11 (15.5)		2 (13.3)	2 (6.1)	
Triple negative	7 (17.1)	11 (15.5)		1 (6.7)	5 (15.2)	
Histological grade			0.995 <sup>d</sup>			0.748 <sup>d</sup>
Low	1 (2.4%)	3 (4.2%)		0 (0.0%)	2 (6.1%)	
Moderate	28 (68.3%)	46 (64.8%)		13 (86.7%)	26 (78.8%)	
High	12 (29.3%)	22 (31.0%)		2 (13.3%)	5 (15.2%)	
pT category			0.015 <sup>d</sup>			1.000
1	10 (24.4)	32 (45.1)		5 (33.3)	11 (33.3)	
2	26 (63.4)	36 (50.7)		10 (66.7)	22 (66.7)	
3	5 (12.2)	3 (4.2)		0 (0)	0 (0)	

**Table 1** (continued)

Characteristics	Training dataset (n = 112)		p value	Validation dataset (n = 48)		p value
	LVI present (n = 41)	LVI absent (n = 71)		LVI present (n = 15)	LVI absent (n = 33)	
pN category			0.000 <sup>d</sup>			0.003 <sup>d</sup>
0	8 (19.5)	50 (70.4)		5 (33.3)	24 (72.7)	
1	16 (39.0)	13 (18.3)		3 (20.0)	4 (12.1)	
2	7 (17.1)	4 (5.6)		2 (13.3)	4 (12.1)	
3	10 (24.4)	4 (5.6)		5 (33.3)	1 (3.0)	
Stages			0.000 <sup>d</sup>			0.037 <sup>d</sup>
1	4 (9.8)	28 (39.4)		2 (13.3)	9 (27.3)	
2	18 (43.9)	34 (47.9)		6 (40.0)	19 (57.6)	
3	19 (46.3)	9 (12.7)		7 (46.7)	5 (15.2)	
Type of surgery			0.004			0.136 <sup>c</sup>
Mastectomy	38 (92.7)	49 (69.0)		14 (93.3)	23 (69.7)	
BCS	3 (7.3)	22 (31.0)		1 (6.7)	10 (30.3)	
Adjuvant chemotherapy			1.000 <sup>c</sup>			0.642 <sup>c</sup>
No	2 (4.9)	3 (4.2)		2 (13.3)	3 (9.1)	
Yes	39 (95.1)	68 (95.8)		13 (86.7)	30 (90.9)	
Adjuvant radiotherapy			0.387			0.269
No	25 (61.0)	49 (69.0)		7 (46.7)	21 (63.6)	
Yes	16 (39.0)	22 (31.0)		8 (53.3)	12 (36.4)	
Adjuvant endocrine therapy			0.652			1.000 <sup>c</sup>
No	15 (36.6)	23 (32.4)		4 (26.7)	9 (27.3)	
Yes	26 (63.4)	48 (67.6)		11 (73.3)	24 (72.7)	

*Abbreviations:* LVI, lymphovascular invasion; NME, nonmass enhancement; *mrALN status*, MRI axillary lymph nodes status; AVS, adjacent vessel sign; BCS, breast-conserving surgery; TIC, time-intensity curve

<sup>a</sup>Data are medians, with interquartile ranges in parentheses

<sup>b</sup>p values were obtained with the Mann–Whitney U test

<sup>c</sup>p values were obtained with the Fisher exact test

<sup>d</sup>p values were obtained with the  $\chi^2$  test for trend. Unless otherwise specified, continuous variables were compared by the Student t-test and categorical variables were compared using the  $\chi^2$  test

**Table 2** Multivariate logistic analysis of the risk factors for LVI

Variable	Clinico-radiological model		Radiomics model	
	Odds ratio	p value	Odds ratio	p value
Peritumoral edema	3.500 (1.611–7.604)	0.002	3.439 (1.134–11.178)	0.032
AVS	2.266 (1.025–5.010)	0.043	NA	NA
<i>mrALN status</i>	3.503 (1.625–7.552)	0.001	NA	NA
Rad-score	NA	NA	2.666 (1.862–4.256)	<0.001

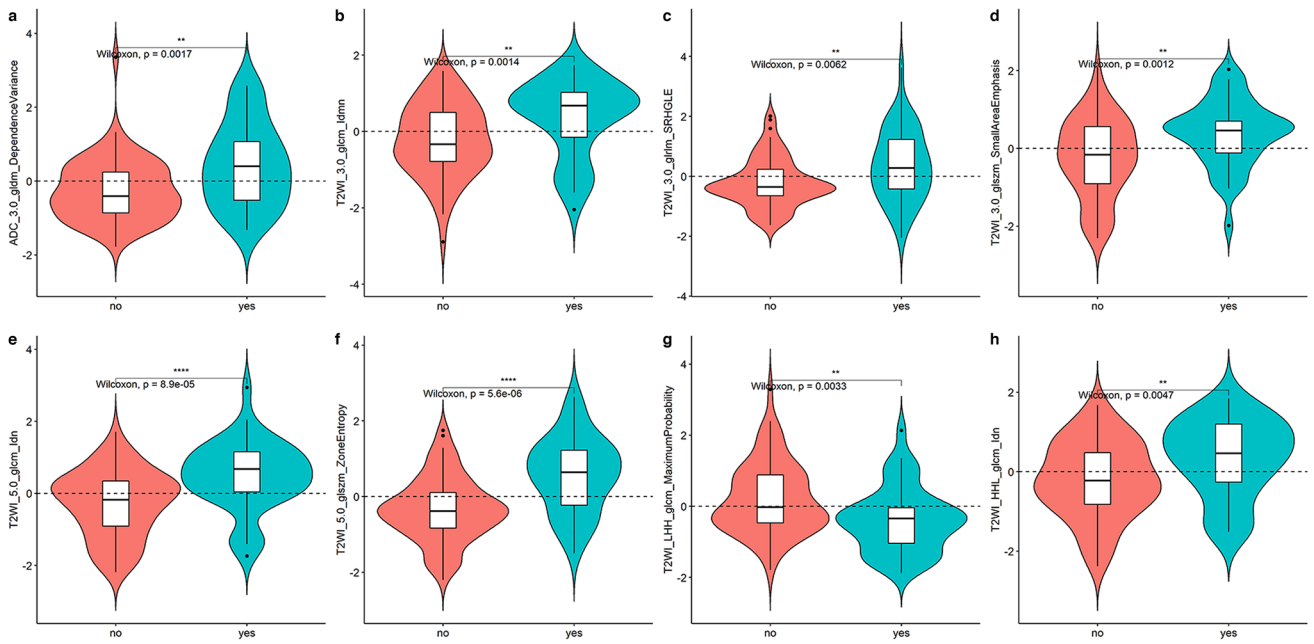
*Abbreviations:* AVS, adjacent vessel sign; *mrALN status*, MRI axillary lymph nodes status

analysis of the predictors of DFS in the entire cohort. The results of the multivariate Cox regression analysis showed that the Rad-score was an independent predictor of DFS (HR, 4.535; 95%CI: 1.428–14.404,  $p = 0.010$ ), following the peritumoral edema (Table 4).

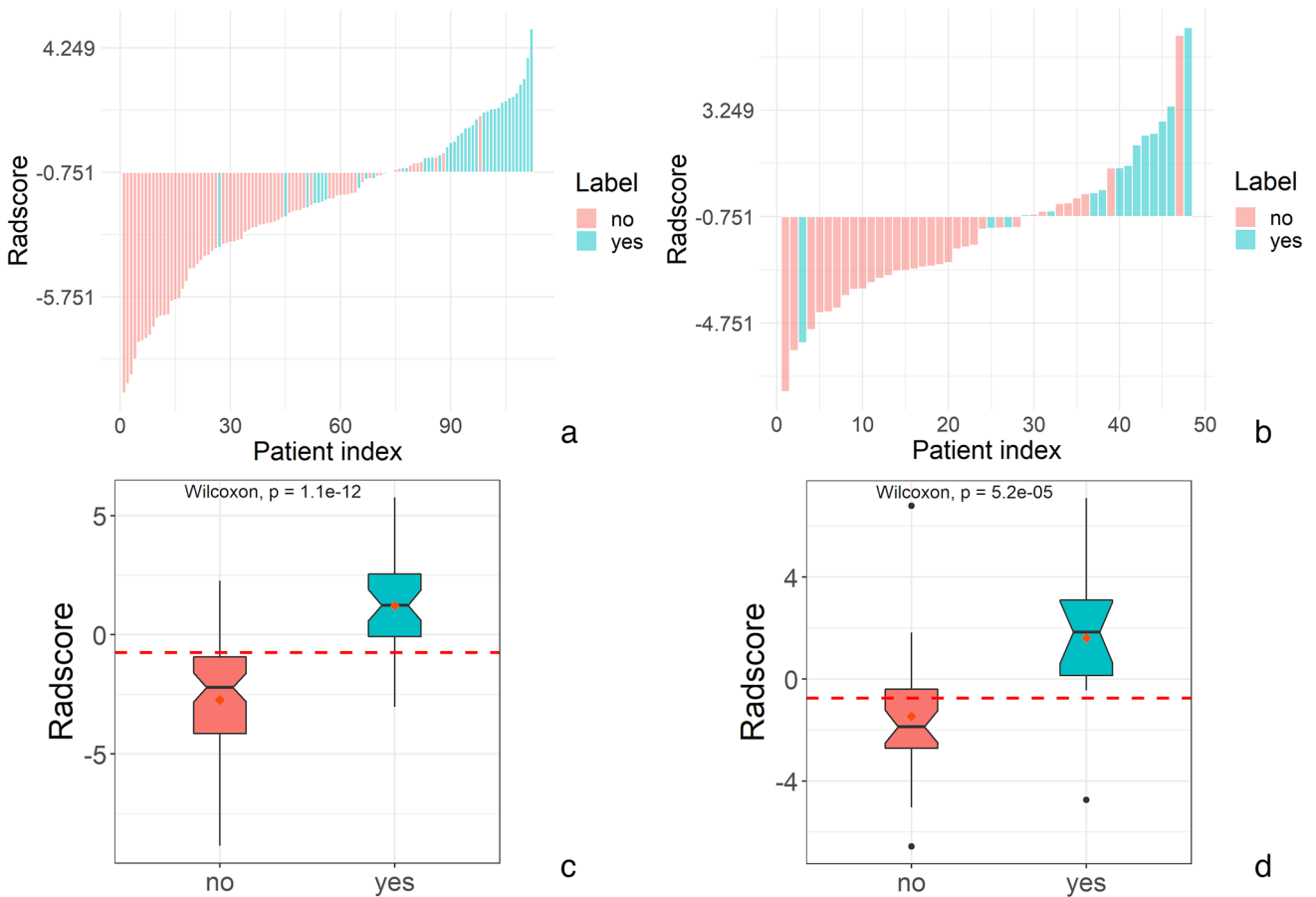
## Discussion

In the present study, we developed and validated a radiomics signature derived from multiparametric MRI, which was capable of predicting LVI in patients with breast IDC.



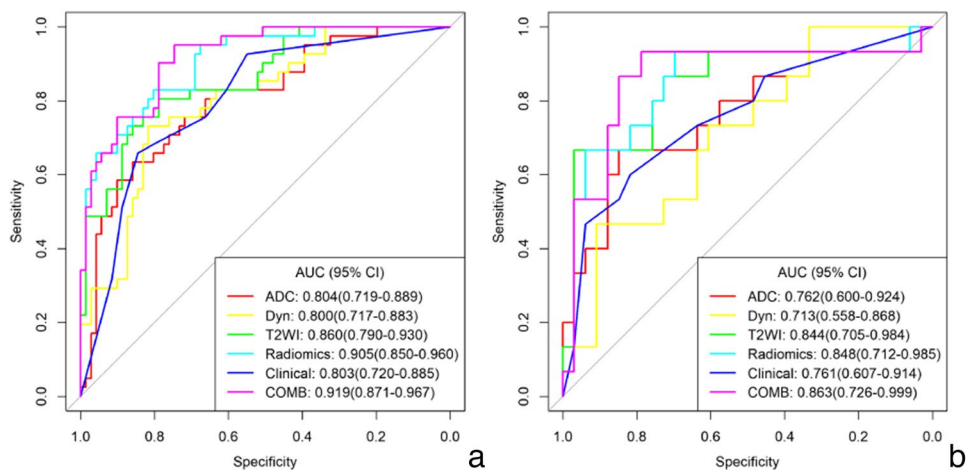


**Fig. 1** Plots (a–h) illustrate the eight radiomics features with a significant difference between the LVI-positive and LVI-negative groups in the training dataset



**Fig. 2** Plots (a, b) show the Rad-score for each patient, and plots (c, d) depict the Rad-score in the training and validation datasets, respectively

**Fig. 3** Plots (a, b) display the ROC curves of the radiomics signatures, the clinico-radiological model, and the radiomics nomogram for LVI status in the training and validation datasets, respectively



**Table 3** Diagnostic performance of the radiomics signatures, clinico-radiological model, and the radiomics nomogram for LVI status

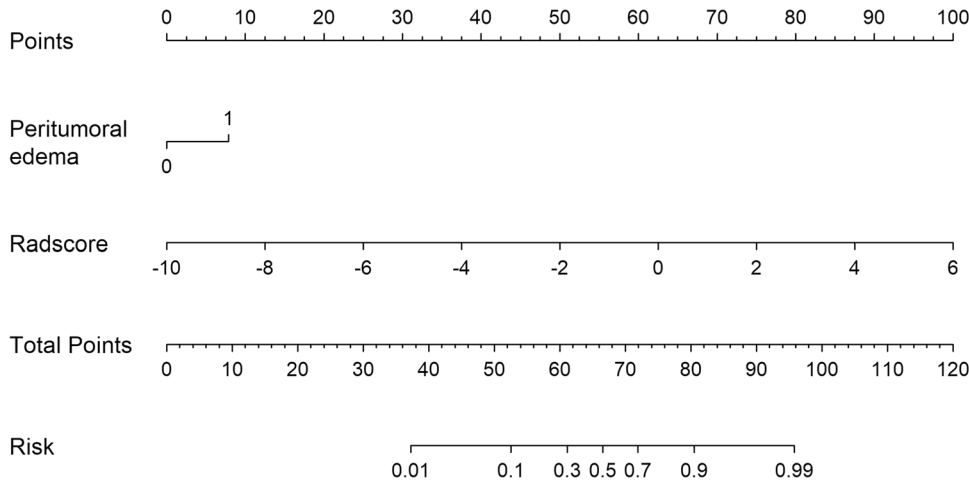
Model	C-index(95%CI)	
	Training dataset	Validation dataset
Radiomics signature-ADC	0.804 (0.719–0.889)	0.762 (0.600–0.924)
Radiomics signature-cT1WI	0.800 (0.717–0.883)	0.713 (0.558–0.868)
Radiomics signature-T <sub>2</sub> WI	0.860 (0.790–0.930)	0.844 (0.705–0.984)
Radiomics signature-fusion	0.905 (0.850–0.960)	0.848 (0.712–0.985)
Clinico-radiological model	0.803 (0.720–0.885)	0.761 (0.607–0.914)
Radiomics nomogram	0.919 (0.871–0.967)	0.863 (0.726–0.999)

providing a straightforward and noninvasive approach for personalized prediction of LVI preoperatively. Additionally, the nomogram was significantly associated with DFS, playing an important role in predicting the prognosis in patients with breast IDC.

A recent study considered that DCE-MRI-based radiomics signature was an independent risk factor for predicting LVI status of IDC patients [21]. In their study, the researchers only evaluated the clinical value of radiomics derived from contrast-enhanced MRI. While in the present study, multiparametric MRI-based radiomics features, including T2WI, ADC maps, and cT1WI, were extracted and used for LVI prediction. The results demonstrated that the fusion radiomics signature from the combination of all the sequences achieved a better predictive efficacy than either of them alone for LVI status of breast cancer. Different sequences can reflect different characteristics of tumors, such as tumor intensity, histologic cellularity, and microvessel density. Therefore, multiparametric MRI features can make full use of the advantages of each sequence to reflect more comprehensive tumor information. Additionally, seven

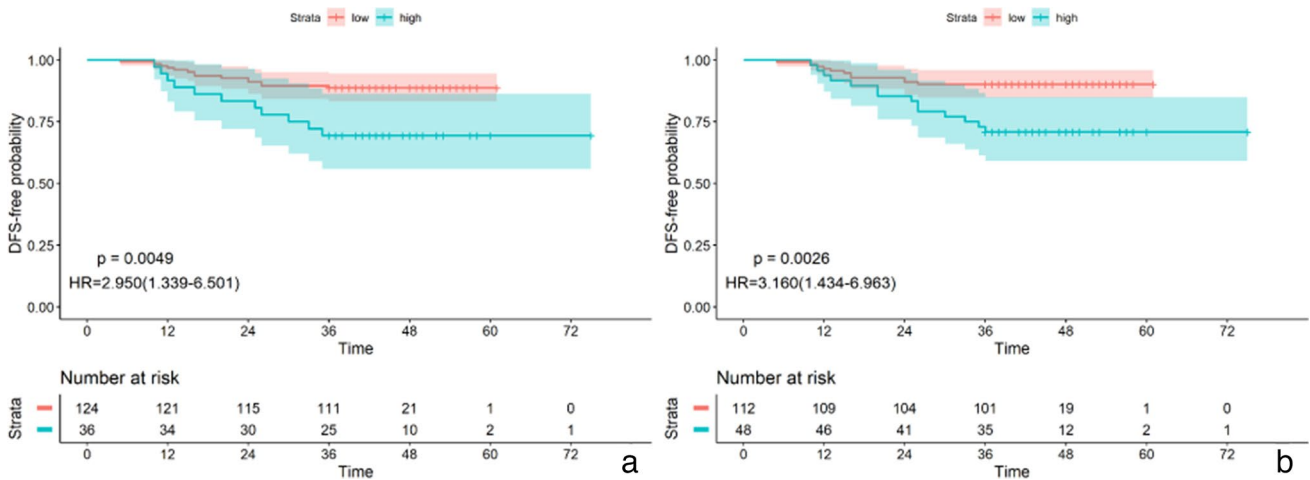
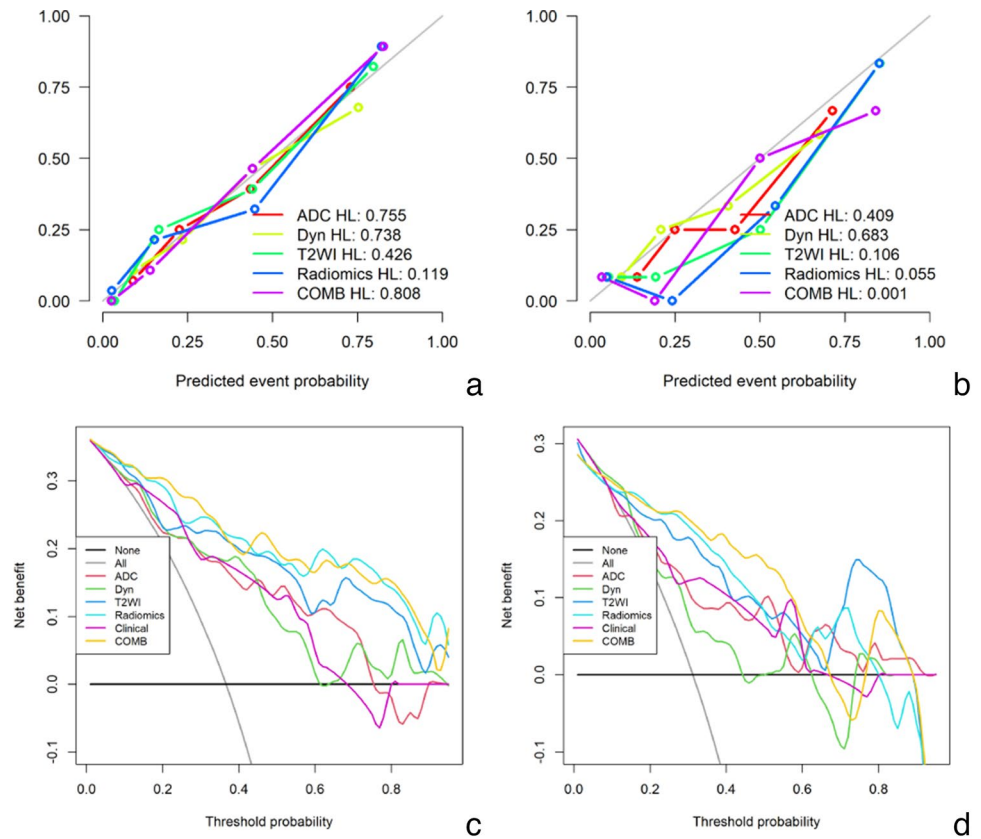
Furthermore, the proposed radiomics nomogram that integrated MRI features and the radiomics signature exhibited a high accuracy for predicting LVI preoperatively,

**Fig. 4** The developed radiomics nomogram for the prediction of LVI status in the training dataset





**Fig. 5** Plots (a, b) illustrate the calibration curves of the radiomics nomogram in the training and validation datasets, respectively. Plots (c, d) show the results of DCA for the radiomics nomogram in the training and validation datasets, respectively. The Y-axis represents the net benefit. The X-axis indicates the threshold probability



**Fig. 6** Kaplan–Meier survival analysis according to the radiomics signature (a) and the nomogram (b) in the entire cohort

of the eight selected features were derived from T2-weighted images, while the DCE-MRI-based radiomics signature was not included. This is consistent with previous speculation that T2-weighted images capture better tumor heterogeneity than contrast-enhanced T1-weighted images [26, 27].

For the fusion radiomics signature, eight radiomics features were selected after the multivariate logistic regression analysis, including two wavelet features and six LoG features. The LoG filter employs a Gaussian filter to smooth

the images, which may help decrease the influence of noise. It can also enhance textural details, which will help improve the efficiency of capturing phenotypic features mapped to tumoral heterogeneity [28, 29]. The wavelet transformation can help us quantify high-dimensional multi-frequency tumor information that are difficult to explain intuitively [30]. Wavelet features mirror multi-frequency information of the tumor in a different wavelet-transformed frequency range [30]. LoG and wavelet features are extracted from

**Table 4** Multivariable Cox regression analysis of predictors of DFS in the entire cohort

Variable	Hazard ratio(95% CI)	<i>p</i> value
Peritumoral edema		
Negative	Ref	
Positive	3.062 (0.875–10.711)	0.080
Stage		
1	Ref	
2	2.406 (0.260–22.223)	0.439
3	19.703 (2.160–179.700)	0.008
Adjuvant radiotherapy		
No	Ref	
Yes	3.066 (1.051–8.945)	0.040
Rad-score	4.535 (1.428–14.404)	0.010

multiple frequency and spatial domain, which are difficult to explain intuitively in clinic. Nevertheless, through radiomics analysis, these high-dimensional image features were detected and successfully used in the prediction of efficacy of neoadjuvant therapy, lymph node metastasis, and gene expression [28–30]. In our study, all these high-dimensional features were significantly higher in the LVI-positive group than those in the LVI-negative group.

In addition, peritumoral edema, AVS, and mrALN status were independent predictors of LVI in the clinico-radiological model in this study. However, only peritumoral edema showed a sufficient predictive power with the adjunction of radiomics signature into the multivariable logistic regression model. Then, peritumoral edema was incorporated to develop a radiomics nomogram model, which showed a satisfactory discrimination ability for predicting LVI, with an AUC value of 0.863 (95% CI: 0.726–0.999) in the validation dataset. The biological mechanism of the association between peritumoral edema and LVI remains elusive. Peritumoral edema may be caused by the increased vascular permeability and peritumoral injection of cytokines [10, 24]. Uematsu et al. pointed out that edema may be caused by obstruction and dilation of lymphatic drainage massage by tumor emboli, and a high degree of LVI is significantly associated with edema [12, 13].

Histological LVI has been reported to be associated with a poorer prognosis of breast cancer in numerous studies [4, 31–33]. Additionally, there is a significant difference in survival outcomes between the high- and low-risk groups defined by our radiomics-based models, in which patients with high-risk LVI had a significantly poorer DFS, compared with those with low-risk LVI. As the outcomes in the LVI high-risk group were not satisfactory postoperatively, neoadjuvant therapy may be a better choice for patients in the high-risk group to obtain drug sensitivity in vivo and to kill potential metastatic cells as early as possible. Therefore,

based on our radiomics model, it is feasible to guide the treatment plan and implement personalized treatment.

Several limitations of this study should be pointed out. Firstly, this was a retrospective single-center study, indicating the necessity to a large-scale prospective multicenter study to further validate the effectiveness of the proposed radiomics nomogram. Secondly, multiphase DCE-MRI is commonly used in breast contrast-enhanced scan of breast, while only the early images after enhancement were used to extract radiomics features in this study. Finally, the DCE-MRI sequence of our institution was originally designed for the analysis of quantitative pharmacokinetic parameters. The spatial resolution was slightly inferior to the conventional contrast T1W sequence, which may affect the robustness of these radiomics features. Moreover, extracting radiomics features of quantitative parameter (k<sub>trance</sub>, k<sub>ep</sub>, and V<sub>e</sub>) of DCE-MRI may provide more information for identification of LVI status, so that further studies are needed to verify this aspect as well.

In conclusion, radiomics features derived from multiparametric MRI are potential biomarkers for predicting LVI. The proposed radiomics nomogram, incorporating multiparametric MRI-based radiomics signature and MR-reported peritumoral edema, exhibited a promising predictive efficacy for LVI status and clinical outcomes, and may provide valuable guidance for individualized treatment in patients with breast IDC.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00330-021-08504-6>.

**Funding** This study was supported by the National Natural Science Foundation of China (No. 82171923, 82001789 and 81802479), the Applied Basic Research Programs of Shanxi Province (No.201801D121307 and 201801D221390), the Key Research and Development (R&D) Projects of Shanxi Province (No. 201803D31168), the Youth Project of Shanxi Provincial Health Commission (No. 2019058), and the Open Fund from Shanxi Medical University-Collaborative Innovation Center for Molecular Imaging of Precision Medicine (No. 2020-MS01).

## Declarations

**Guarantor** The scientific guarantor of this publication is Xiaotang Yang.

**Conflict of interest** One of the authors (JR) is an employee of GE Healthcare. The remaining authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

**Statistics and biometry** No complex statistical methods were necessary for this paper.

**Informed consent** Written informed consent was waived by the Institutional Review Board.

**Ethical approval** Institutional Review Board approval was obtained.

### Methodology

- retrospective
- diagnostic or prognostic study
- performed at one institution

### References

- Sung H, Ferlay J, Siegel RL et al (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 71:209–249
- Weigelt B, Peterse JL, van 't Veer LJ (2005) Breast cancer metastasis: markers and models. *Nat Rev Cancer* 5:591–602
- Kurozumi S, Joseph C, Sonbul S et al (2019) A key genomic subtype associated with lymphovascular invasion in invasive breast cancer. *Br J Cancer* 120:1129–1136
- Rakha EA, Martin S, Lee AH et al (2012) The prognostic significance of lymphovascular invasion in invasive breast carcinoma. *Cancer* 118:3670–3680
- Cheung SM, Husain E, Mallikourti V, Masannat Y, Heys S, He J (2021) Intra-tumoural lipid composition and lymphovascular invasion in breast cancer via non-invasive magnetic resonance spectroscopy. *Eur Radiol* 31:3703–3711
- Zhou P, Jin C, Lu J et al (2021) The value of nomograms in preoperative prediction of lymphovascular invasion in primary breast cancer undergoing modified radical surgery: based on multiparametric ultrasound and clinicopathologic indicators. *Ultrasound Med Biol* 47:517–526
- Vasconcelos I, Hussainzada A, Berger S et al (2016) The St. Gallen surrogate classification for breast cancer subtypes successfully predicts tumor presenting features, nodal involvement, recurrence patterns and disease free survival. *Breast* 29:181–185
- Shen S, Wu G, Xiao G et al (2018) Prediction model of lymphovascular invasion based on clinicopathological factors in Chinese patients with invasive breast cancer. *Medicine (Baltimore)* 97:e12973
- Mann RM, Cho N, Moy L (2019) Breast MRI: state of the art. *Radiology* 292:520–536
- Cheon H, Kim HJ, Lee SM et al (2017) Preoperative MRI features associated with lymphovascular invasion in node-negative invasive breast cancer: a propensity-matched analysis. *J Magn Reson Imaging* 46:1037–1044
- Igarashi T, Furube H, Ashida H, Ojiri H (2018) Breast MRI for prediction of lymphovascular invasion in breast cancer patients with clinically negative axillary lymph nodes. *Eur J Radiol* 107:111–118
- Uematsu T (2015) Focal breast edema associated with malignancy on T2-weighted images of breast MRI: peritumoral edema, prepectoral edema, and subcutaneous edema. *Breast Cancer* 22:66–70
- Uematsu T, Kasami M, Watanabe J (2014) Is evaluation of the presence of prepectoral edema on T2-weighted with fat-suppression 3 T breast MRI a simple and readily available noninvasive technique for estimation of prognosis in patients with breast cancer? *Breast Cancer* 21:684–692
- Mori N, Mugikura S, Takasawa C et al (2016) Peritumoral apparent diffusion coefficients for prediction of lymphovascular invasion in clinically node-negative invasive breast cancer. *Eur Radiol* 26:331–339
- Lambin P, Leijenaar RTH, Deist TM et al (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 14:749–762
- Kiessling F (2018) The changing face of cancer diagnosis: from computational image analysis to systems biology. *Eur Radiol* 28:3160–3164
- Xu X, Zhang HL, Liu QP et al (2019) Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol* 70:1133–1144
- Yang L, Gu D, Wei J et al (2019) A radiomics nomogram for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Liver Cancer* 8:373–386
- Nie P, Yang G, Wang N et al (2021) Additional value of metabolic parameters to PET/CT-based radiomics nomogram in predicting lymphovascular invasion and outcome in lung adenocarcinoma. *Eur J Nucl Med Mol Imaging* 48:217–230
- Luo Y, Mei D, Gong J, Zuo M, Guo X (2020) Multiparametric MRI-based radiomics nomogram for predicting lymphovascular space invasion in endometrial carcinoma. *J Magn Reson Imaging* 52:1257–1262
- Liu Z, Feng B, Li C et al (2019) Preoperative prediction of lymphovascular invasion in invasive breast cancer with dynamic contrast-enhanced-MRI-based radiomics. *J Magn Reson Imaging* 50:847–857
- Wolff AC, Hammond MEH, Allison KH et al (2018) Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *J Clin Oncol* 36:2105–2122
- Goldhirsch A, Winer EP, Coates AS et al (2013) Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* 24:2206–2223
- Cheon H, Kim HJ, Kim TH et al (2018) Invasive breast cancer: prognostic value of peritumoral edema identified at preoperative MR imaging. *Radiology* 287:68–75
- Hyun SJ, Kim EK, Moon HJ, Yoon JH, Kim MJ (2016) Preoperative axillary lymph node evaluation in breast cancer patients by breast magnetic resonance imaging (MRI): can breast MRI exclude advanced nodal disease? *Eur Radiol* 26:3865–3873
- Park H, Lim Y, Ko ES et al (2018) Radiomics signature on magnetic resonance imaging: association with disease-free survival in patients with invasive breast cancer. *Clin Cancer Res* 24:4705–4714
- Parikh J, Selmi M, Charles-Edwards G et al (2014) Changes in primary breast cancer heterogeneity may augment midtreatment MR imaging assessment of response to neoadjuvant chemotherapy. *Radiology* 272:100–112
- Zhou X, Yi Y, Liu Z et al (2019) Radiomics-based pretherapeutic prediction of non-response to neoadjuvant therapy in locally advanced rectal cancer. *Ann Surg Oncol* 26:1676–1684
- Cui Y, Liu H, Ren J et al (2020) Development and validation of a MRI-based radiomics signature for prediction of KRAS mutation in rectal cancer. *Eur Radiol* 30:1948–1958
- Wu Q, Wang S, Chen X et al (2019) Radiomics analysis of magnetic resonance imaging improves diagnostic performance of lymph node metastasis in patients with cervical cancer. *Radiother Oncol* 138:141–148
- Liu YL, Saraf A, Lee SM et al (2016) Lymphovascular invasion is an independent predictor of survival in breast cancer after neoadjuvant chemotherapy. *Breast Cancer Res Treat* 157:555–564
- Hamy AS, Lam GT, Laas E et al (2018) Lymphovascular invasion after neoadjuvant chemotherapy is strongly associated with poor prognosis in breast carcinoma. *Breast Cancer Res Treat* 169:295–304
- Zhang S, Zhang D, Gong M, Wen L, Liao C, Zou L (2017) High lymphatic vessel density and presence of lymphovascular invasion both predict poor prognosis in breast cancer. *BMC Cancer* 17:335

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