Breast Cancer Research

RESEARCH

Open Access

Automatic segmentation-based multi-modal radiomics analysis of US and MRI for predicting disease-free survival of breast cancer: a multicenter study

Lang Xiong[1†], Xiaofeng Tang[2†], Xinhua Jiang[3†], Haolin Chen[4], Binyan Qian[5], Biyun Chen[3], Xiaofeng Lin[3], Jianhua Zhou[2*] and Li Li[3*]

Abstract

Background Several studies have confirmed the potential value of applying radiomics to predict prognosis of breast cancer. However, the tumor segmentation in these studies depended on delineation or annotation of breast cancer by radiologist, which is often laborious, tedious, and vulnerable to inter- and intra-observer variability. Automatic segmentation is expected to overcome this difficulty. Herein, we aim to investigate the value of automatic segmentation-based multi-modal radiomics signature and magnetic resonance imaging (MRI) features in predicting disease-free survival (DFS) of patients diagnosed with invasive breast cancer.

Methods This retrospective multicenter study included a total of 643 female patients with invasive breast cancer who underwent preoperative ultrasound (US) and MRI for prognostic analysis. Data ($n=480$) from center 1 were divided into training and internal testing sets, while data ($n=163$) from centers 2 and 3 were analyzed as the external testing set. We developed automatic segmentation frameworks for tumor segmentation by deep learning. Then, Least absolute shrinkage and selection operator Cox regression was used to select features to construct radiomics signature, and corresponding radiomics score (Rad-score) was calculated. Finally, six models for predicting DFS were constructed by using Cox regression and assessed in terms of discrimination, calibration, and clinical usefulness.

Results The multi-modal radiomics signature combining intra- and peri-tumoral radiomics signatures of US and MRI achieved a higher C-index in the internal (0.734) and external (0.708) testing sets than most other radiomics signatures in predicting DFS, and successfully stratified patients into low- and high-risk groups. The multi-modal clinical imaging model combining the multi-modal Rad-score and clinical traditional MRI model-score resulted in a higher C-index (0.795) than other models in the external testing set, and it had a better calibration and higher clinical benefit.

[†]Lang Xiong, Xiaofeng Tang and Xinhua Jiang contributed equally to this work.

*Correspondence:
Jianhua Zhou
zhoujh@sysucc.org.cn
Li Li
li2@mail.sysu.edu.cn

Full list of author information is available at the end of the article

Conclusions This study demonstrates that the multi-modal radiomics signature derived from automatic segmentations of US and MRI is a promising risk stratification biomarker for breast cancer, and highlights that the appropriate combination of multi-modal radiomics signature, clinical characteristics, and MRI feature can improve performance of individualized DFS prediction, which might assist in guiding decision-making related to breast cancer.

Keywords Breast cancer, Ultrasound, Magnetic resonance imaging, Multi-modal radiomics, Disease-free survival

Background

Breast cancer is the leading cause of cancer deaths among women worldwide [1]. Despite improvement in breast cancer treatment, a 20.5% recurrence rate within 10 years has been reported [2]. There is a significant difference in the 10-year survival of breast cancer, with 82% among patients without recurrence, compared to 61%, 41%, and 20% in patients with local, regional, or distant recurrence, respectively [3]. Accurate treatment improves high-risk breast cancer outcome [4–6]. Therefore, accurate prediction of recurrence risk can effectively guide clinical decision-making related to breast cancer, thereby improving prognosis of breast cancer.

Radiomics aims to extract quantitative features to develop classifier models for clinical decision-making [7]. Previous studies have applied magnetic resonance imaging (MRI) radiomics [8–11] or ultrasound (US) radiomics [12–14] to predict breast cancer survival. Park et al. successfully developed radiomics signature from MRI as an independent biomarker to predict breast cancer disease-free survival (DFS) [8]. Chitalia et al. identified radiomics phenotypes from MRI and demonstrated its independent and additional value in predicting recurrence-free survival for breast cancer [10]. Yu et al. [12] effectively combined US radiomics signature and clinicopathological features to improve individualized DFS estimation of triple-negative breast cancer (TNBC). Furthermore, we developed US radiomics signature to effectively improve accuracy in predicting DFS for breast cancer [14]. In clinical practice, physicians prefer to make clinical decisions by multi-modal images, as they could provide more supplementary information than that provided by single-modality images. Studies proved multi-modal radiomics outperformed single-modal radiomics for tumor classification and prediction including breast cancer diagnosis [15–17]. Therefore, we want to explore the potential value of combining US and MRI radiomics to predict prognosis of breast cancer in this study.

However, tumor segmentation in most previous radiomics studies depended on delineation or annotation of breast cancer by radiologist, which is often laborious, tedious, and vulnerable to inter- and intra-observer variability [18]. Automatic segmentation is expected to overcome this difficulty in tumor segmentation [19]. Deep learning-based automatic segmentation model not only exhibits excellent robustness against undesired variation, but also demonstrates comparable performance to radiologists for breast tumor segmentation [20]. Qiao et al. developed a tumor automatic segmentation framework (ASF) to improve the diagnostic accuracy and efficiency for breast tumors [21]. Ma et al. [22, 23] and Gan et al. [24] developed radiomics models based on automatically segmented MRI images to perform classification and prediction tasks of breast cancer. To my best knowledge, automatic segmentation-based radiomics analysis of combining US and MRI for predicting breast cancer prognosis has not been reported.

Additionally, many studies have demonstrated that biological changes in the peritumoral tissue of breast tumor might be predictive or prognostic risk factors, such as axillary lymph node metastasis [25, 26], lymphovascular invasion [27], peritumoral edema [28], and neoadjuvant chemotherapy effect [29]. In breast cancer, peritumoral radiomics features could provide complementary information to intratumoral regions in predicting Ki-67 level that is an important prognostic factor [30], and radiomics signature consisting of intra- and peri-tumoral features enabled non-invasive assessment of tumor immune microenvironment that plays pivotal roles in tumor progression, metastasis, and therapeutic response [31]. Based on these findings, the combination of intra- and peri-tumoral features provided important implications for breast cancer prognosis. Thus, whether automatic segmentation-based multi-modal radiomics combining intra- and peri-tumoral features can improve accuracy of breast cancer prognosis prediction merits further investigation.

Herein, we developed an automatic segmentation-based multi-modal radiomics signature combining intra- and peri-tumoral radiomics signatures of US and MRI to improve performance of risk stratification for breast cancer. Furthermore, a multi-modal clinical imaging model combining the multi-modal radiomics score (Rad-score) and clinical traditional MRI model-score was developed to improve performance for individualized DFS prediction.

Methods

Study design and patients

The patient recruitment and design of this study are presented in Fig. 1. First, data from 620 patients undergoing preoperative US at Sun Yat-sen University Cancer Center (SYUCC) (these patients were previously reported [14]), were collected and divided into training and testing sets

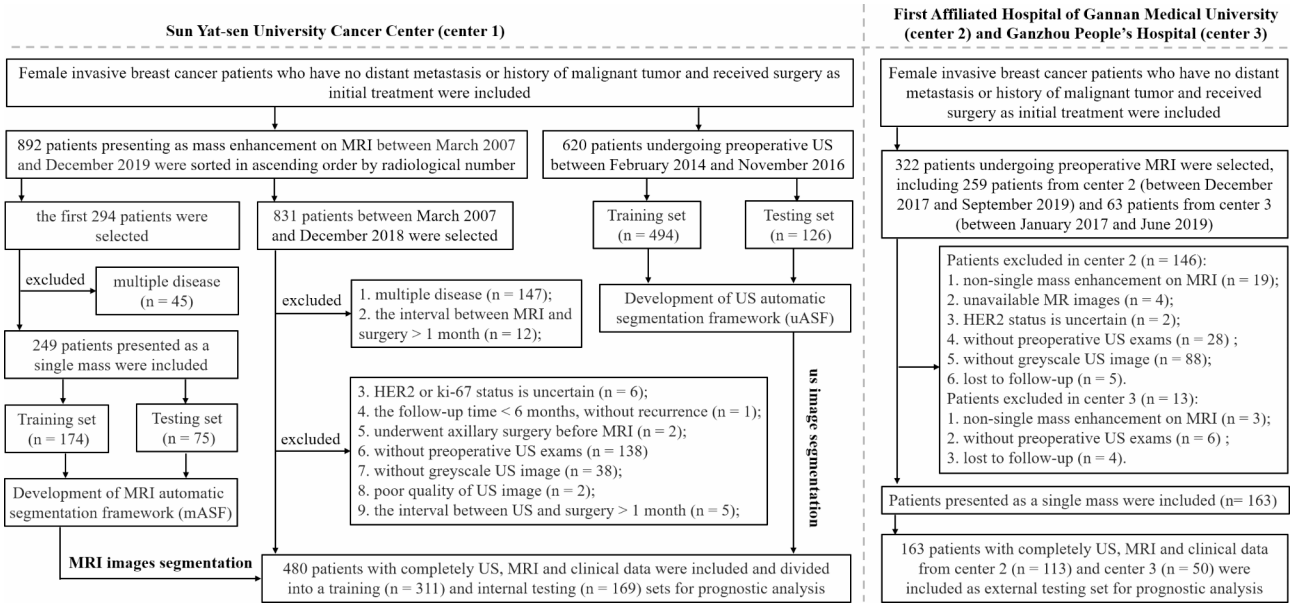

Fig. 1 The patient recruitment and design of this study

to develop and validate the US ASF (uASF), with 26 of them being included in subsequent prognostic analysis; and data from 249 patients undergoing preoperative MRI at SYUCC were collected and divided into training and testing sets to develop and validate the MRI ASF (mASF), with 98 of them being included in subsequent prognostic analysis. Then, data from 643 patients who underwent preoperative US and presented as a single mass on dynamic contrast enhanced (DCE) MRI were collected for prognostic analysis; data from 480 of these patients (mean age, 47.77 years; range, 23–80 years) from SYUCC (between March 2007 and December 2018) were randomly divided into training (n=311) and internal testing (n=169) sets; data from the remaining patients (n=163; mean age, 49.90 years; range, 30–77 years) from the First Affiliated Hospital of Gannan Medical University and Ganzhou People's Hospital were included as external testing set.

Clinical characteristics and follow-up

The Age; menopausal status; breast cancer risk factors (family history of breast cancer and/or history of breast surgery for begin disease); surgery type and postoperative treatment; pathological tumor size; histological type and grade; stages of T, N, and TNM; status of lymphovascular invasion, associated ductal carcinoma in situ; and status of hormone receptor (HR) (including estrogen receptor [ER] and progesterone receptor [PR]), human epidermal growth factor receptor 2 (HER2), and Ki-67 were collected. In immunohistochemistry, HR was considered as positive when nuclear staining presented in at least 1% of tumor cells; HER2 as negative when scored as 0 and 1+, and positive when scored as 3+, while score 2+needed further confirmation with fluorescence in situ hybridization; Ki-67 was considered as high and low expressions when proliferation was ≥14% and <14%, respectively. Tumors were categorized into luminal A, luminal B, HER2-enriched, and TNBC [32].

Study end point, DFS, was defined as time from date of surgery to that of locoregional recurrence, distant metastasis, contralateral breast cancer, or death, whichever came first, and patients without events at the last follow-up were censored.

Image acquisition and segmentation

First, the single US greyscale image (JPEG format) containing the largest section of tumor, and the peak phase of DCE-MRI (DICOM format) according to the time-signal intensity curve (TIC) were selected. Second, the US tumor regions of interests (ROIs) based on manual segmentation (denoted as US-MSeg-intra-ROIs) were used to train a U-Net for developing uASF; the MRI breast mask and MRI tumor ROIs based on manual segmentation (denoted as MRI-MSeg-intra-ROIs) were used to train a three-dimensional (3D) ResUNet and WNet (a network we proposed in this study based on the structure of U-Net) for breast segmentation and breast cancer segmentation respectively (supplementary Figure S1), for developing mASF. The dice similarity coefficient (DSC) of the testing set was calculated to assess segmentation performance. Details regarding image acquisition including DCE-MR protocol (supplementary Table S1), preprocessing and segmentation are provided in the supplement.

Afterwards, images from the 643 patients were inputted into ASFs to generate automatic segmentation-based tumor ROIs (denoted as ASeg-intra-ROIs), which were

then manually checked and adjusted by reader 1 (LX), with these adjustments further validated by reader 2 (XHJ). Both readers collaborated to reach a consensus, thereby ensuring standardization and consistency across different cases. The criteria used for manually checking and adjusting the ASeg-intra-ROIs were as follows: (1) if it only predicted breast cancer region, nothing would be done for MRI-ASeg-intra-ROI. But under the same condition, the missed part of breast cancer in the US-ASeg-intra-ROI would be manually filled, which is due to the fact that it is relatively easy to fill in the missed part for the single US image; (2) if it predicted breast cancer and non-breast cancer regions, the latter would be manually deleted; (3) and if it did not predict breast cancer, it would be replaced by MSeg-intra-ROI. Then, dilated segmentation was performed by morphological expansion of 1–7 pixel for US and 1–7 mm for MRI based on the adjusted ASeg-intra-ROI by using python, and the ASeg-intra-ROI was then subtracted to generate automatic segmentation-based peritumoral ROI (denoted as ASeg-peri-ROI). Altogether, sixteen ROIs were defined for each patient, including one US-ASeg-intra-, seven US-ASeg-peri-, one MRI-ASeg-intra-, and seven MRI-ASeg-peri-ROIs.

Radiomics signature construction

PyRadiomics package of Python, which conforms to the Image Biomarker Standardization Initiative (IBSI) guidelines for radiomic analysis [33], was used to extract radiomics features. Spearman correlation coefficients and Ward linkage method, along with the least absolute shrinkage and selection operator (LASSO) Cox regression model, were performed to select features to construct intra- and peri-radiomics signatures. And corresponding Rad-score was calculated for each patient via a linear combination of the selected features that were weighted by their respective LASSO coefficients. Then, the intra- and peri-radiomics signatures were combined to construct the gross-radiomics signature by using cox regression model. Details regarding radiomics signature construction are provided in the supplement.

Radiomics signature validation

The potential association of radiomics signature with DFS was assessed and validated in the training, internal, and external testing sets. Using the optimal cutoff of Rad-score identified by X-tile [34], patients with Rad-score < cutoff and Rad-score ≥ cutoff were divided into low- and high-risk groups, respectively. Kaplan–Meier survival analysis was subsequently performed to evaluate the survival rates and generate the corresponding survival curves for patients in these two groups. Following this, the log-rank test was applied to determine whether the differences in survival curves between the low- and high-risk groups were statistically significant. Furthermore, Harrell's concordance index (C-index) [35] and time-dependent receiver operating characteristic (ROC) curves were used to evaluate performance of the radiomics signature in predicting DFS. Finally, the DeLong test was utilized to statistically compare the differences between area under the receiver operating characteristic curves (AUC), thereby assessing the discriminatory power of these radiomics signatures.

MRI features

Using the 2013 Breast Imaging Reporting and Data System MRI lexicon, two radiologists (LX and XHJ, with 8 and more than 10 years of experience in breast MRI interpretation, respectively), independently reviewed MRI features and reached a joint decision by consensus. These features evaluated in this study included MR tumor size (the largest diameter of tumor); shape (oval, round, or irregular); margin (circumscribed, irregular, or spiculated); internal enhancement (homogeneous, heterogeneous, or rim enhancement); and TIC (persistent, plateau, or washout types), which was calculated by drawing an ROI on the fastest-enhancing area of tumor. Three months after the first review, 60 patients were randomly selected and reviewed again to evaluate intra-observer agreement of categorical and continuous features by using the Kappa test and intraclass correlation coefficient (ICC).

Model construction and evaluation

In order to investigate the value of multi-modal radiomics signature and MRI features for DFS prediction, we constructed six predictive models by using cox regression model, including a clinical model, a traditional MRI model, a clinical traditional MRI model, a multi-modal radiomics signature, a multi-modal radiomics model, and a multi-modal clinical imaging model.

We first used the univariate Cox regression model to analyze the relationship of clinical and MRI features with DFS. Then, we used multivariate Cox proportional hazards model to select the best combination of clinical predictors by including variables in a step-wise manner based on the Bayesian information criterion (BIC), to build the clinical model. And the same method was used to select the best combination of MRI predictors to build the traditional MRI model. Subsequently, we combined the selected clinical and MRI predictors to build the clinical traditional MRI model. We also combined the US-intra-, US-peri-, MRI-intra-, and MRI-peri-radiomics signatures to construct the multi-modal radiomics signature, which was then incorporated into the clinical model to build the multi-modal radiomics model. Finally, we combined the multi-modal Rad-score and the clinical traditional MRI model-score to construct the multi-modal

clinical imaging model. The multi-modal Rad-score and the clinical traditional MRI model-score were calculated by the score formulation (score = ∑ feature values × coefficient of feature) using the features selected for constructing the multi-modal radiomics signature and the clinical traditional MRI model, respectively.

These developed models were comprehensively evaluated and compared, focusing on the following four terms: (1) the C-index, which measures the agreement between the model predicted DFS and the actual DFS observed in all patients, was calculated to assess the discrimination power of models; (2) calibration curves were plotted to evaluate the consistency between predicted survival probabilities and actual survival probabilities at different time points for models; (3) decision curve analysis (DCA) was conducted to evaluate the clinical usefulness of models by calculating net benefits across different threshold probabilities [36]; and (4) the BIC values were calculated to evaluate the goodness-of-fit of models.

Statistical analysis

Python 3.7.11 and R 4.0.3 were used for statistical analyses. A bilateral $P<0.05$ was considered as significant. Categorical and continuous variables were compared using the chi-squared test and Kruskal–Wallis H-test, respectively. The python package "lifelines" was used to perform the Kaplan–Meier survival analysis, log-rank test, and Cox regression. The R packages "glment" and "survival" were used to perform the LASSO-Cox regression analysis. The R packages "rms" and "rmda" were used to draw curves of calibration and DCA, respectively.

Results

Patient characteristics

Table 1 shows the patient characteristics and MRI features in the three datasets. Based on twice reviews of MRI features for the 60 selected patients, the mean ICC of MRI tumor size was 0.991, and the Kappa values for shape, margin, internal enhancement, and TIC type were 0.864, 0.612, 0.856, and 0.737, respectively. The median follow-up time was 52.00 (interquartile range, 39.93–74.30) months. During follow-up, 74 positive events, including 17 locoregional recurrences, 35 distant metastases, 14 concurrently locoregional and distant metastases, three contralateral breast cancer, and five deaths occurred.

Segmentation

The mean DSC of uASF and mASF for the testing set were 0.91 and 0.73, respectively. The segmentation examples are shown in Figs. 2 and 3. For US-ASeg-intra-ROIs, 54.43% had no adjustment, 45.10% had non-breast cancer region removed or missed parts of breast cancer region filled, and 0.47% were replaced by US-MSeg-intra-ROIs. For MRI-ASeg-intra-ROIs, 67.03% had no adjustment, 15.86% had non-breast cancer region deleted, and 17.11% were replaced by MRI-MSeg-intra-ROIs.

Radiomics signature construction and validation

The developed intra- and peri-radiomics signatures are shown in supplementary Tables S2 and S3. Using the Rad-score cutoff, the US-peri (3 pixel) and MRI-peri (2, 3 and 4)-radiomics signatures successfully stratified patients into high- and low-risk groups in all the three datasets ($P<0.05$), whereas the other intra- and peri-radiomics signatures did not (supplementary Figures S2 and S3).

Table 2 shows that the US-gross (6 pixel)-radiomics signature combining US-intra- and peri (6 pixel)-radiomics signatures and the MRI-gross (2 mm)-radiomics signature combining MRI-intra- and peri (2 mm)-radiomics signatures achieved higher C-index (0.722 and 0.684) than the other US and MRI radiomics signatures in the internal testing set, respectively. Thus, we combined the US-gross (6 pixel)- and MRI-gross (2 mm)-radiomics signatures to construct multi-modal radiomics signature, which achieved higher C-index (0.734) than the other radiomics signatures in the internal testing set and higher C-index (0.708) than most radiomics signatures in the external testing set. It also successfully stratified patients into high- and low-risk groups in all the three datasets based on a cutoff value of 1.36 (Fig. 4, $P<0.05$).

Table 3 and supplementary Figure S4 show higher AUCs of the US-gross (6 pixel)-radiomics signature in predicting 1-, 3-, and 5-year DFS (0.786, 0.753, and 0.774) than those of MRI-gross (2 mm)-radiomics signature (0.685, 0.674, and 0.730) in the internal testing set, as well as higher AUCs of the US-gross (6 pixel)-radiomics signature in predicting 1- and 3-year DFS (0.820 and 0.741) than those of MRI-gross (2 mm)-radiomics signature (0.648 and 0.714) in the external testing set. Compared to the US-gross (6 pixel)- and MRI-gross (2 mm)-radiomics signatures, the multi-modal radiomics signature achieved higher AUCs (0.762 and 0.792) in predicting 3- and 5-year DFS for the internal testing set and higher AUCs (0.835, 0.797, and 0.658) in predicting 1-, 3-, and 5-year DFS for the external testing set, although the difference was not significant ($P>0.05$).

Model construction and evaluation

Table S4 indicates that T stage, N stage, TNM stage, and MR tumor size were associated with DFS in the univariate analysis. Table 4 displays that the clinical model included N stage and histological grade; the traditional MRI model included MR tumor size; the clinical traditional MRI model included N stage, histological grade, and MR tumor size; the multi-modal radiomics signature included US-intra-, US-peri (6 pixel)-, MRI-intra-, and

Table 1 Clinical characteristics and magnetic resonance imaging features of patients

Characteristics	Training set (*n* = 311)	Internal testing set (*n* = 169)	External testing set (*n* = 163)	*P* value
Age, years*	47.32 [23.00–76.00]	48.59 [27.00–80.00]	49.90 [30.00–77.00]	0.040
Menopausal status				0.062
Pre-menopause	218 (70.10%)	109 (64.50%)	97 (59.51%)	
Post-menopause	93 (29.90%)	60 (35.50%)	66 (40.49%)	
Risk factors				0.037
No	285 (91.64%)	155 (91.72%)	159 (97.55%)	
Yes	26 (8.36%)	14 (8.28%)	4 (2.45%)	
Pathologic tumor size (cm) *	2.09 [0.30–7.00]	2.27 [0.40–6.60]	2.35 [0.60–6.00]	0.008
ER				0.998
Negative	67 (21.54%)	36 (21.30%)	35 (21.47%)	
Positive	244 (78.46%)	133 (78.70%)	128 (78.53%)	
PR				0.825
Negative	93 (29.90%)	46 (27.22%)	47 (28.83%)	
Positive	218 (70.10%)	123 (72.78%)	116 (71.17%)	
HER2				0.927
Negative	232 (74.60%)	128 (75.74%)	124 (76.07%)	
Positive	79 (25.40%)	41 (24.26%)	39 (23.93%)	
Ki-67				0.353
Low	81 (26.05%)	43 (25.44%)	33 (20.25%)	
High	230 (73.95%)	126 (74.56%)	130 (79.75%)	
Molecular subtype				0.761
Luminal A	69 (22.19%)	39 (23.08%)	30 (18.40%)	
Luminal B	180 (57.88%)	98 (57.99%)	99 (60.74%)	
HER2-enriched	27 (8.68%)	9 (5.33%)	13 (7.98%)	
Triple-negative	35 (11.25%)	23 (13.61%)	21 (12.88%)	
T stage				0.563
1	183 (58.84%)	94 (55.62%)	88 (53.99%)	
2	120 (38.59%)	71 (42.01%)	72 (44.17%)	
3 or 4	8 (2.57%)	4 (2.37%)	3 (1.84%)	
N stage				0.456
0	193 (62.06%)	113 (66.86%)	101 (61.96%)	
1	76 (24.44%)	37 (21.89%)	42 (25.77%)	
2	29 (9.32%)	14 (8.28%)	9 (5.52%)	
3	13 (4.18%)	5 (2.96%)	11 (6.75%)	
TNM stage				0.627
I	133 (42.77%)	64 (37.87%)	62 (38.04%)	
II	133 (42.77%)	84 (49.70%)	78 (47.85%)	
III	45 (14.47%)	21 (12.43%)	23 (14.11%)	
Lymphovascular invasion				0.022
Absent	211 (67.85%)	119 (70.41%)	130 (79.75%)	
Present	100 (32.15%)	50 (29.59%)	33 (20.25%)	
Associated ductal carcinoma in situ				< 0.0001
Absent	200 (64.31%)	118 (69.82%)	136 (83.44%)	
Present	111 (35.69%)	51 (30.18%)	27 (16.56%)	
Histological type				0.584
invasive ductal carcinoma	284 (91.32%)	154 (91.12%)	154 (94.48%)	
invasive lobular carcinoma	12 (3.86%)	6 (3.55%)	2 (1.23%)	
Others invasive breast cancer	15 (4.82%)	9 (5.33%)	7 (4.29%)	
Histological grade				0.0002
I or II	175 (56.27%)	94 (55.62%)	123 (75.46%)	
III	112 (36.01%)	60 (35.50%)	27 (16.56%)	
Missing	24 (7.72%)	15 (8.88%)	13 (7.98%)	

Table 1 (continued)

Characteristics	Training set (*n* = 311)	Internal testing set (*n* = 169)	External testing set (*n* = 163)	*P* value
Surgery type				0.795
Mastectomy	192 (61.74%)	100 (59.17%)	102 (62.58%)	
Breast conservation	119 (38.26%)	69 (40.83%)	61 (37.42%)	
Adjuvant endocrine therapy				0.828
No	85 (27.33%)	42 (24.82%)	42 (25.77%)	
Yes	226 (72.67%)	127 (75.15%)	121 (74.23%)	
Adjuvant chemotherapy				0.0002
No	72 (23.15%)	44 (26.04%)	15 (9.20%)	
Yes	239 (76.85%)	125 (73.96%)	148 (90.80%)	
Adjuvant radiotherapy				0.823
No	149 (47.91%)	86 (50.89%)	80 (49.08%)	
Yes	162 (52.09%)	83 (49.11%)	83 (50.92%)	
Adjuvant targeted therapy				0.096
No	256 (82.32%)	146 (86.39%)	126 (77.30%)	
Yes	55 (17.68%)	23 (13.61%)	37 (22.70%)	
MR tumor size (cm) *	2.18 [0.70–6.00]	2.36 [0.70–7.20]	2.31 [0.90–5.90]	0.119
Time-signal intensity curve				0.132
Persistent	17 (5.47%)	4 (2.37%)	6 (3.68%)	
Plateau	181 (58.20%)	106 (62.72%)	112 (68.71%)	
Washout	113 (36.33%)	59 (34.91%)	45 (27.61%)	
Shape				0.008
Oval or Round	86 (27.65%)	55 (32.54%)	29 (17.79%)	
Irregular	225 (72.35%)	114 (67.46%)	134 (82.21%)	
Margin				0.717
Circumscribed	36 (11.58%)	21 (12.43%)	24 (14.72%)	
Irregular	154 (49.52%)	77 (45.56%)	71 (43.56%)	
Spiculated	121 (38.91%)	71 (42.01%)	68 (41.72%)	
Internal enhancement				0.061
Homogeneous	6 (1.93%)	4 (2.37%)	2 (1.23%)	
Heterogeneous	218 (70.10%)	108 (63.91%)	125 (76.69%)	
Rim enhancement	87 (27.97%)	57 (33.73%)	36 (22.09%)	
Disease-free survival events				0.974
No	275 (88.42%)	149 (88.17%)	145 (88.96%)	
Yes	36 (11.58%)	20 (11.83%)	18 (11.04%)	
Follow-up time (months) *	61.48 [1.57-138.27]	61.36 [5.47–138.10]	47.63 [7.97–75.97]	< 0.0001

*Data were represented by mean [minimum, maximum]. ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; MR, magnetic resonance; TNM, tumor-node-metastasis

MRI-peri (2 mm)-radiomics signatures; the multi-modal radiomics model included multi-modal radiomics signature, N stage, and histological grade; the multi-modal clinical imaging model included multi-modal Rad-score and clinical traditional MRI model-score.

Table 5 shows the performance of models. The clinical model, traditional MRI model, clinical traditional MRI model, multi-modal radiomics signature, multi-modal radiomics model and multi-modal clinical imaging model achieved C-index of 0.628, 0.680, 0.667, 0.734, 0.703, and 0.719 for the internal testing set, and of 0.724, 0.718, 0.777, 0.708, 0.734, and 0.795 for the external testing set, respectively. In terms of goodness-of-fit, the multi-modal clinical imaging model performed best with the lowest BIC (371.342). The multi-modal clinical imaging model achieved better calibration than the clinical traditional MRI model and multi-modal radiomics signature in predicting 1-, 3-, and 5-year DFS for the internal testing set. It also achieved better calibration than the multi-modal radiomics signature in predicting 1- and 3-year DFS for the external testing set (Fig. 5).

Within the probability threshold range of 0.00–0.45, the multi-modal clinical imaging model had a higher overall net benefit than the clinical traditional MRI model and multi-modal radiomics signature in the external testing set (supplementary Figure S5); this finding was confirmed when the internal and external testing sets were merged into a testing set, and when the training, internal

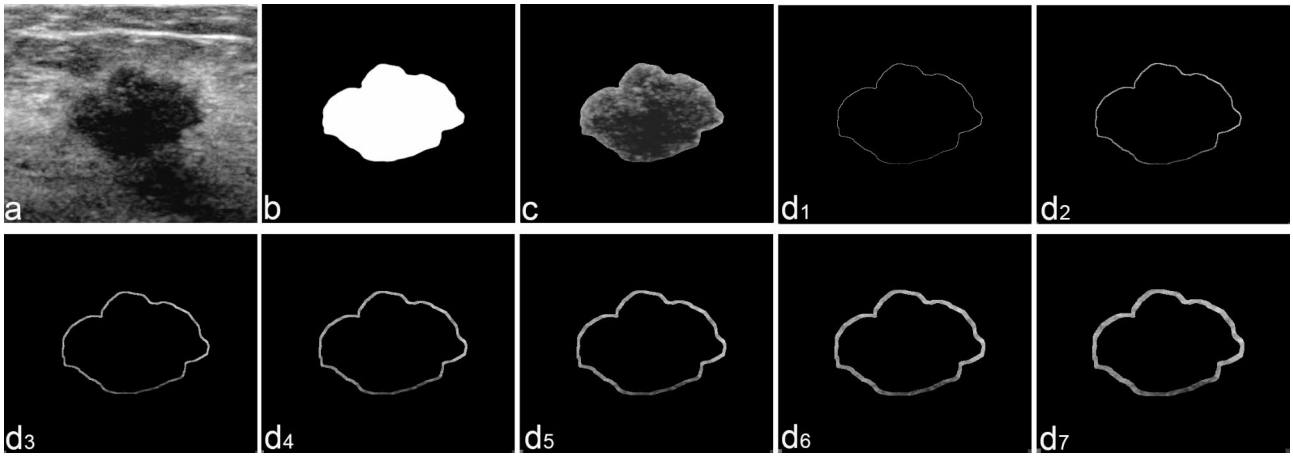

Fig. 2 The ultrasound automatic segmentation example: **a** Original image of breast cancer. **b** Tumor mask. **c** Tumor region in the ultrasound image. **d** Peritumoral tissue with different thicknesses (1–7 pixel) in the ultrasound image

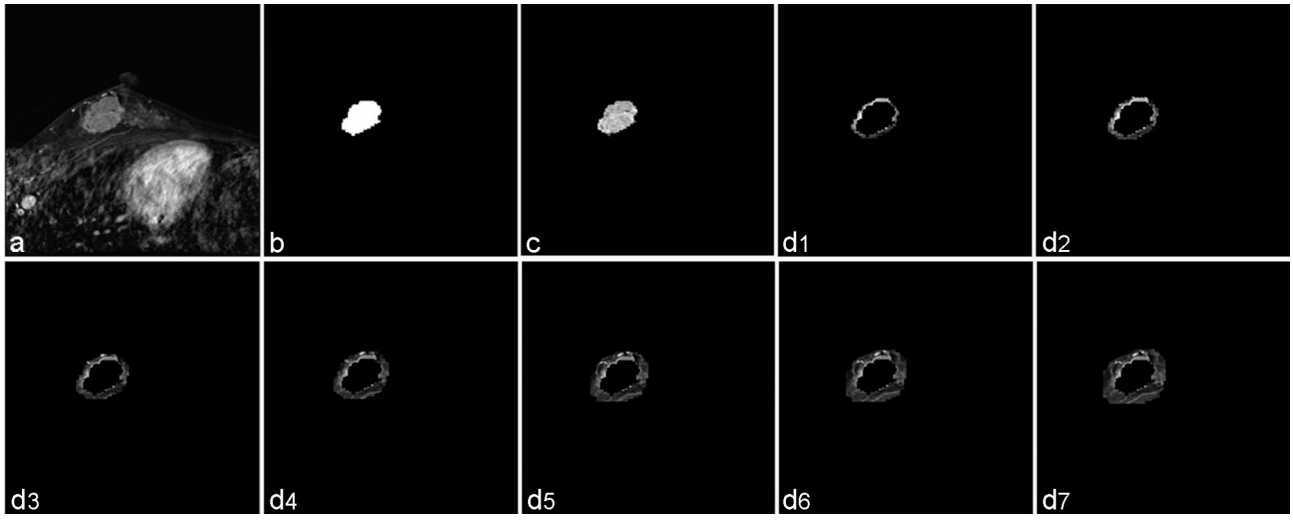

Fig. 3 The MRI segmentation example: **a** Original image of breast cancer. **b** Tumor mask. **c** Tumor region in the MRI image. **d** Peritumoral tissue with different thicknesses (1–7 mm) in the MRI image. MRI, magnetic resonance imaging

and external testing sets were merged into a whole dataset (Fig. 6).

Discussion

In this study, we developed an automatic segmentation-based multi-modal radiomics signature to effectively perform risk stratification for invasive breast cancer. The multi-modal clinical imaging model combining the automatic segmentation-based multi-modal Rad-score and clinical traditional MRI model-score performed best among all the models, demonstrating the incremental value of the automatic segmentation-based multi-modal radiomics signature and MRI tumor size in predicting DFS for patients with invasive breast cancer.

Both ASFs demonstrated promising segmentation performances, with the uASF achieving a higher DSC (0.91) than that reported in previous studies [37] and the mASF achieving a similar or slightly lower DSC (0.73) than that reported in related studies [20, 38]. Zhang et al. trained a 3D framework using images of the pre-contrast, first post-contrast, and subtracted sequences collected from the GE and Siemens systems and reported a DSC of 0.72 for a testing set consisting of 48 cases [38]. Recently, Rahimpour et al. trained a 3D U-Net by using the first post-contrast and subtraction images collected from the GE and Siemens systems and achieved a DSC of 0.78 for a testing set consisting of 30 cases [20]. While it is difficult to directly compare studies performed on different datasets, our current ASFs may be more robust because they were trained and tested using images from a wider range of sources.

The automatic segmentation-based radiomics signatures were significantly associated with DFS and patients with higher Rad-scores experienced worse outcomes.

Table 2 C-index of radiomics signature in disease-free survival prediction

Radiomics signature	C-index (95% confidence interval)		
	Training set	Internal testing set	External testing set
US-intra-radiomics signature	0.707 (0.608–0.807)	0.675 (0.539–0.811)	0.645 (0.501–0.790)
US-peri (1 pixel)-radiomics signature	0.720 (0.622–0.819)	0.650 (0.513–0.787)	0.647 (0.502–0.791)
US-peri (2 pixel)-radiomics signature	0.802 (0.713–0.891)	0.676 (0.541–0.812)	0.597 (0.451–0.742)
US-peri (3 pixel)-radiomics signature	0.848 (0.767–0.929)	0.610 (0.472–0.749)	0.639 (0.494–0.783)
US-peri (4 pixel)-radiomics signature	0.812 (0.725–0.900)	0.625 (0.487–0.763)	0.642 (0.497–0.787)
US-peri (5 pixel)-radiomics signature	0.761 (0.667–0.855)	0.702 (0.568–0.835)	0.553 (0.408–0.698)
US-peri (6 pixel)-radiomics signature	0.760 (0.665–0.854)	0.717 (0.585–0.850)	0.646 (0.502–0.791)
US-peri (7 pixel)-radiomics signature	0.729 (0.631–0.826)	0.685 (0.550–0.820)	0.596 (0.451–0.742)
US-gross (1 pixel)-radiomics signature	0.741 (0.644–0.837)	0.682 (0.547–0.817)	0.656 (0.512–0.800)
US-gross (2 pixel)-radiomics signature	0.804 (0.716–0.893)	0.689 (0.554–0.824)	0.597 (0.451–0.742)
US-gross (3 pixel)-radiomics signature	0.849 (0.769–0.930)	0.618 (0.480–0.756)	0.640 (0.495–0.784)
US-gross (4 pixel)-radiomics signature	0.816 (0.730–0.903)	0.639 (0.502–0.777)	0.644 (0.499–0.788)
US-gross (5 pixel)-radiomics signature	0.776 (0.684–0.869)	0.709 (0.575–0.842)	0.582 (0.437–0.727)
US-gross (6 pixel)-radiomics signature	0.774 (0.682–0.867)	0.722 (0.590–0.854)	0.654 (0.510–0.798)
US-gross (7 pixel)-radiomics signature	0.748 (0.652–0.844)	0.694 (0.560–0.829)	0.609 (0.463–0.754)
MRI-intra-radiomics signature	0.651 (0.549–0.753)	0.636 (0.498–0.774)	0.696 (0.555–0.838)
MRI-peri (1 mm)-radiomics signature	0.759 (0.664–0.854)	0.639 (0.501–0.777)	0.588 (0.442–0.733)
MRI-peri (2 mm)-radiomics signature	0.688 (0.588–0.789)	0.670 (0.534–0.806)	0.651 (0.506–0.795)
MRI-peri (3 mm)-radiomics signature	0.681 (0.580–0.782)	0.654 (0.517–0.791)	0.719 (0.581–0.858)
MRI-peri (4 mm)-radiomics signature	0.685 (0.584–0.786)	0.653 (0.516–0.790)	0.729 (0.591–0.867)
MRI-peri (5 mm)-radiomics signature	0.681 (0.581–0.782)	0.660 (0.523–0.797)	0.584 (0.439–0.730)
MRI-peri (6 mm)-radiomics signature	0.676 (0.575–0.777)	0.677 (0.541–0.813)	0.557 (0.412–0.702)
MRI-peri (7 mm)-radiomics signature	0.680 (0.579–0.781)	0.674 (0.538–0.810)	0.567 (0.422–0.712)
MRI-gross (1 mm)-radiomics signature	0.658 (0.556–0.760)	0.635 (0.497–0.773)	0.696 (0.554–0.837)
MRI-gross (2 mm)-radiomics signature	0.699 (0.599–0.799)	0.684 (0.549–0.819)	0.684 (0.542–0.827)
MRI-gross (3 mm)-radiomics signature	0.685 (0.584–0.785)	0.665 (0.529–0.802)	0.729 (0.592–0.867)
MRI-gross (4 mm)-radiomics signature	0.680 (0.579–0.780)	0.665 (0.528–0.801)	0.733 (0.596–0.871)
MRI-gross (5 mm)-radiomics signature	0.674 (0.573–0.775)	0.652 (0.515–0.789)	0.693 (0.552–0.835)
MRI-gross (6 mm)-radiomics signature	0.680 (0.579–0.781)	0.665 (0.529–0.802)	0.651 (0.507–0.795)
MRI-gross (7 mm)-radiomics signature	0.679 (0.578–0.779)	0.660 (0.523–0.797)	0.668 (0.524–0.811)
Multi-modal radiomics signature*	0.777 (0.684–0.869)	0.734 (0.603–0.864)	0.708 (0.568–0.848)

* The multi-modal radiomics signature were developed by combining the US-intra-, US-peri (6 pixel)-, MRI-intra- and MRI-peri (2 mm)-radiomics signatures. C-index, Harrel's concordance index; US, ultrasound; MRI, magnetic resonance imaging

For predicting DFS of invasive breast cancer, both the US-gross (6 pixel)- and MRI-gross (2 mm)-radiomics signatures achieved comparable C-index (0.722 and 0.684, respectively) or AUC values to those reported in previous non-automatic segmentation based radiomics studies [8, 12, 14, 39]. We developed an US-radiomics signature with C-index of 0.632 [14], and a recent study developed an US-radiomics signature with C-index of 0.70 and 0.69 in the internal and external validation cohorts [12]. Park et al. developed an MRI-radiomics signature with C-index of 0.67 [8]. Yu et al. developed an MRI radiomics signature with an AUC of 0.66 in 3-year DFS prediction [39], which was similar to that reported for the MRI-gross (2 mm)-radiomics signature (0.674). These findings suggested that the radiomics signature used automatic segmentation method can achieve comparable performance to those reported in previous studies that used manual or semi-automatic segmentation methods, thereby confirming the feasibility of automatic segmentation-based radiomics analysis in predicting breast cancer survival. Moreover, the combination of intra- and peritumoral radiomics signatures improved accuracy of DFS prediction for invasive breast cancer.

The US-gross (6 pixel)-radiomics signature achieved a higher C-index and higher AUC than the MRI-gross (2 mm)-radiomics signature in predicting DFS; however, whether this was due to the better segmentation performance of uASF remains to be determined. Although the gain of multi-modal radiomics signature combining US-gross (6 pixel)- and MRI-gross (2 mm)-radiomics signatures over single-modality radiomics signatures (US and MRI radiomics signature) was not statistically significant ($P>0.05$), it achieved a higher AUC. Whether this was related to the relatively lower segmentation accuracy for MRI, resulting in a less pronounced gain of the multi-modal radiomics signature over the single-modality

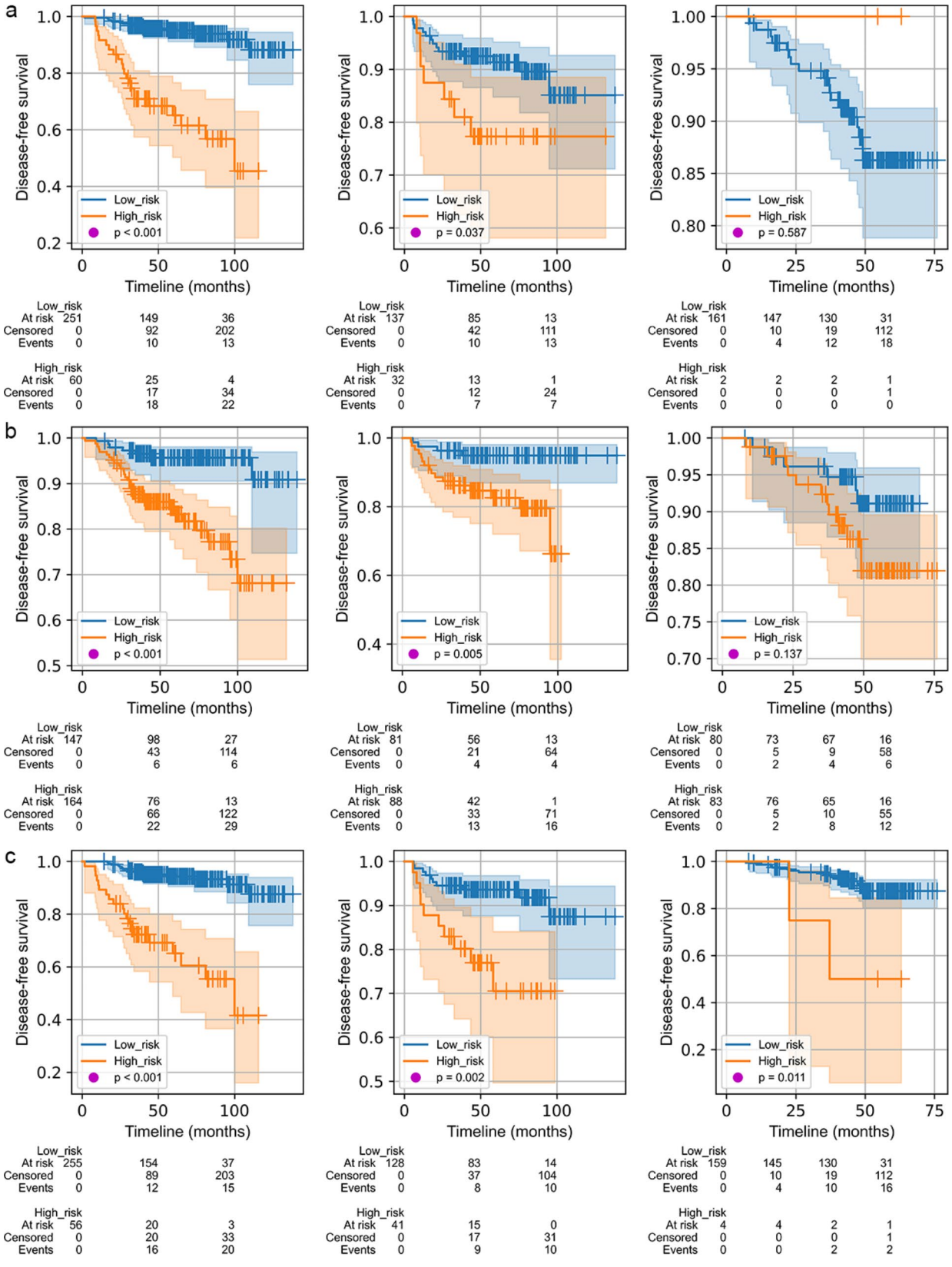

Fig. 4 Kaplan–Meier survival analyses according to the US-gross (6 pixel)-radiomics signature (**a**), the MRI-gross (2 mm)-radiomics signature (**b**) and the multi-modal radiomics signature (**c**), survival curves for the training, internal and external testing sets are shown on the left, middle and right of each subgraph, respectively. A significant association of the multi-modal radiomics signature with the disease-free survival was observed in the training set, which was confirmed in both the internal and external testing sets. US, ultrasound; MRI, magnetic resonance imaging

Table 3 AUC (95% confidence interval) of radiomics signature in disease-free survival prediction

Radiomics signature	Training set			Internal testing set			External testing set		
	1-year	3-year	5-year	1-year	3-year	5-year	1-year	3-year	5-year
US-gross (6 pixel)-radiomics signature	0.916 (0.804–1.000)	0.807 (0.712–0.901)	0.783 (0.687–0.880)	0.786 (0.640–0.933)	0.753 (0.654–0.852)	0.774 (0.666–0.883)	0.820 (0.691–0.949)	0.741 (0.621–0.861)	0.640 (0.470–0.810)
MRI-gross (2 mm)-radiomics signature	0.820 (0.669–0.972)	0.715 (0.615–0.815)	0.747 (0.657–0.836)	0.685 (0.474–0.896)	0.674 (0.557–0.790)	0.730 (0.616–0.843)	0.648 (0.424–0.872)	0.714 (0.579–0.849)	0.647 (0.484–0.810)
Multi-modal radiomics signature	0.948 (0.894–1.000)	0.798 (0.703–0.894)	0.807 (0.721–0.893)	0.765 (0.591–0.939)	0.762 (0.661–0.863)	0.792 (0.683–0.900)	0.835 (0.773–0.897)	0.797 (0.673–0.922)	0.658 (0.490–0.827)
$^{*}P$	0.492	0.859	0.551	0.519	0.717	0.455	0.843	0.088	0.643
$^{**}P$	0.043	0.059	0.151	0.354	0.102	0.182	0.064	0.118	0.877

Delong test was performed between AUCs of the multi-modal radiomics and the other two radiomics signatures in predicting DFS for invasive breast cancer. $^{*}P$ value of the multi-modal radiomics signature compared to the US-gross (6 pixel)-radiomics signature in predicting 1-, 3- and 5-year DFS. $^{**}P$ value of the multi-modal radiomics signature compared to the MRI-gross (2 mm)-radiomics signature in predicting 1-, 3- and 5-year DFS. AUC, Area under receiver operating characteristic curves; US, ultrasound; MRI, magnetic resonance imaging; DFS, disease-free survival

radiomics signatures, requires further investigation. New deep learning approaches and prospective multicenter data are needed in the future to solve this issue. The proposed multi-modal radiomics signature not only outperformed the single-modality radiomics signatures in terms of AUC and C-index values, but also successfully classified patients with different prognoses into low- and high-risk groups. Thus, these findings suggested that the combination of US and MRI radiomics improved performances of risk stratification and DFS prediction for patients with invasive breast cancer.

Furthermore, the clinical traditional MRI model achieved a higher C-index (0.667) than the clinical model (0.628) in the internal testing set, and a higher C-index (0.777) than the clinical model (0.724) and traditional MRI model (0.718) in the external testing set, demonstrating the incremental value of MR tumor size in predicting DFS. The multi-modal radiomics model achieved a higher C-index in the internal (0.703) and external (0.734) testing sets than the clinical model, demonstrating the incremental value of multi-modal radiomics signature in DFS prediction. Note that the multi-modal clinical imaging model achieved a lower C-index (0.719) than the multi-modal radiomics signature (0.734) in the internal testing set—this might be related to tumor heterogeneity. Interestingly, the multi-modal clinical imaging model, which was more parsimonious (lower BIC value), achieved the highest C-index (0.795) in the external testing set, as well as better calibration than multi-modal radiomics signature in predicting 1-, and 3-year DFS for both the internal and external testing sets. When predicting 5-year DFS, the calibration of multi-modal clinical imaging model only outperformed that of multi-modal radiomics signature for the internal testing set, it did not superior for the external testing set. We think this may be due to external factors (e.g. patient demographics, tumor biology, etc.) influencing model's calibration. More importantly, DCA results demonstrated that the multi-modal clinical imaging model outperformed other models, confirming that the appropriate combination of multi-modal radiomics signature, N stage, histological grade, and MRI tumor size further improved DFS predictive performance.

Therefore, we believe that under the premise of protecting patient data privacy, the application of an artificial intelligence-based DFS prediction model—the developed multi-modal clinical imaging model, could assist clinicians in making individualized clinical decisions by providing valuable information in clinical practice and ultimately improving patient prognosis. In addition, radiologists could make a comprehensive interpretation of a patient based on the specific radiomics signature, the clinical and MRI features.

Although, the present radiomics analysis is based on the radiomics features extracted from manual checked and adjusted ASeg-intra-ROIs rather than entirely automatic radiomics features. More than half of the ASeg-intra-ROIs did not require manual adjustment, which could decrease the degree of manual intervention during radiomics analysis, thus potentially reduce the radiologist's workload and the variability in tumor-based radiomics. Moreover, the developed US-gross (6 pixel)- and MRI-gross (2 mm)-radiomics signatures achieved comparable performance to previous manual- or semi-automatic segmentation-based radiomics analysis in predicting DSF of patients with breast cancer, and the developed automatic segmentation-based multi-modal radiomics signature achieved better performance.

Our study had some limitations. First, due to the retrospective nature of the study, there is a potential risk of patient selection bias and uncontrolled scanning parameters, but an independent external testing set was applied and images were preprocessed in order to minimize discrepancies respectively. Second, the MRI features

Table 4 Models constructed based on the results of the multivariate analysis in the training set

Models	Variables	Hazard ratio (95% CI)	*P* value
Clinical model	N stage		
	0	Reference	/
	1	1.36 (0.99–1.89)	0.06
	2	1.04 (0.73–1.50)	0.82
	3	1.46 (1.19–1.78)	< 0.005
	Histological grade		
	I or II	Reference	/
	III	1.38 (0.99–1.93)	0.05
	Missing	0.83 (0.48–1.43)	0.50
Traditional MRI model	MR tumor size (cm)	1.37 (1.04–1.82)	0.03
Clinical traditional MRI model	N stage		
	0	Reference	/
	1	1.33 (0.95–1.84)	0.09
	2	1.01 (0.70–1.46)	0.96
	3	1.43 (1.16–1.75)	< 0.005
	Histological grade		
	I or II	Reference	/
	III	1.31 (0.93–1.86)	0.12
	Missing	0.83 (0.49–1.43)	0.51
	MR tumor size (cm)	1.18 (0.84–1.65)	0.34
Multi–modal radiomics signature	US–intra–radiomics signature	1.48 (1.09–2.01)	0.01
	US–peri (6 pixel)–radiomics signature	2.46 (1.63–3.70)	< 0.005
	MRI–intra–radiomics signature	6.43 (1.60–25.82)	0.01
	MRI–peri (2 mm)–radiomics signature	1.36 (0.95–1.96)	0.10
Multi–modal radiomics model	US–intra–radiomics signature	1.51 (1.08–2.13)	0.02
	US–peri (6 pixel)–radiomics signature	2.44 (1.48–4.02)	< 0.005
	MRI–intra–radiomics signature	49.08 (8.35–288.54)	< 0.005
	MRI–peri (2 mm)–radiomics signature	1.51 (1.05–2.18)	0.03
	N stage		
	0	Reference	/
	1	0.73 (0.48–1.12)	0.15
	2	0.86 (0.56–1.33)	0.50
	3	1.31 (1.08–1.60)	0.01
	Histological grade		
	I or II	Reference	/
	III	0.71 (0.48–1.06)	0.09
	Missing	0.84 (0.44–1.61)	0.61
Multi–modal clinical imaging model	Multi-modal Rad-score	1.31 (1.02–1.69)	0.04
	Clinical traditional MRI model-score	2.22 (1.37–3.58)	< 0.005

CI, confidence interval; US, ultrasound; MRI, magnetic resonance imaging; Rad-score, radiomics score

Table 5 Performance of models in disease-free survival prediction

Models	C-index (95% confidence interval)			BIC value
	Training set	Internal testing set	External testing set	
Clinical model	0.696 (0.596–0.796)	0.628 (0.490–0.766)	0.724 (0.586–0.862)	387.381
Traditional MRI model	0.639 (0.536–0.741)	0.680 (0.544–0.815)	0.718 (0.579–0.857)	385.130
Clinical traditional MRI model	0.727 (0.629–0.824)	0.667 (0.531–0.804)	0.777 (0.647–0.908)	390.107
Multi-modal radiomics signature	0.777 (0.684–0.869)	0.734 (0.603–0.864)	0.708 (0.568–0.848)	414.786
Multi-modal radiomics model	0.766 (0.672–0.859)	0.703 (0.569–0.837)	0.734 (0.597–0.872)	522.804
Multi-modal clinical imaging model	0.781 (0.689–0.873)	0.719 (0.586–0.851)	0.795 (0.667–0.922)	371.342

C-index, Harrel's concordance index; BIC, Bayesian information criterion; MRI, magnetic resonance imaging

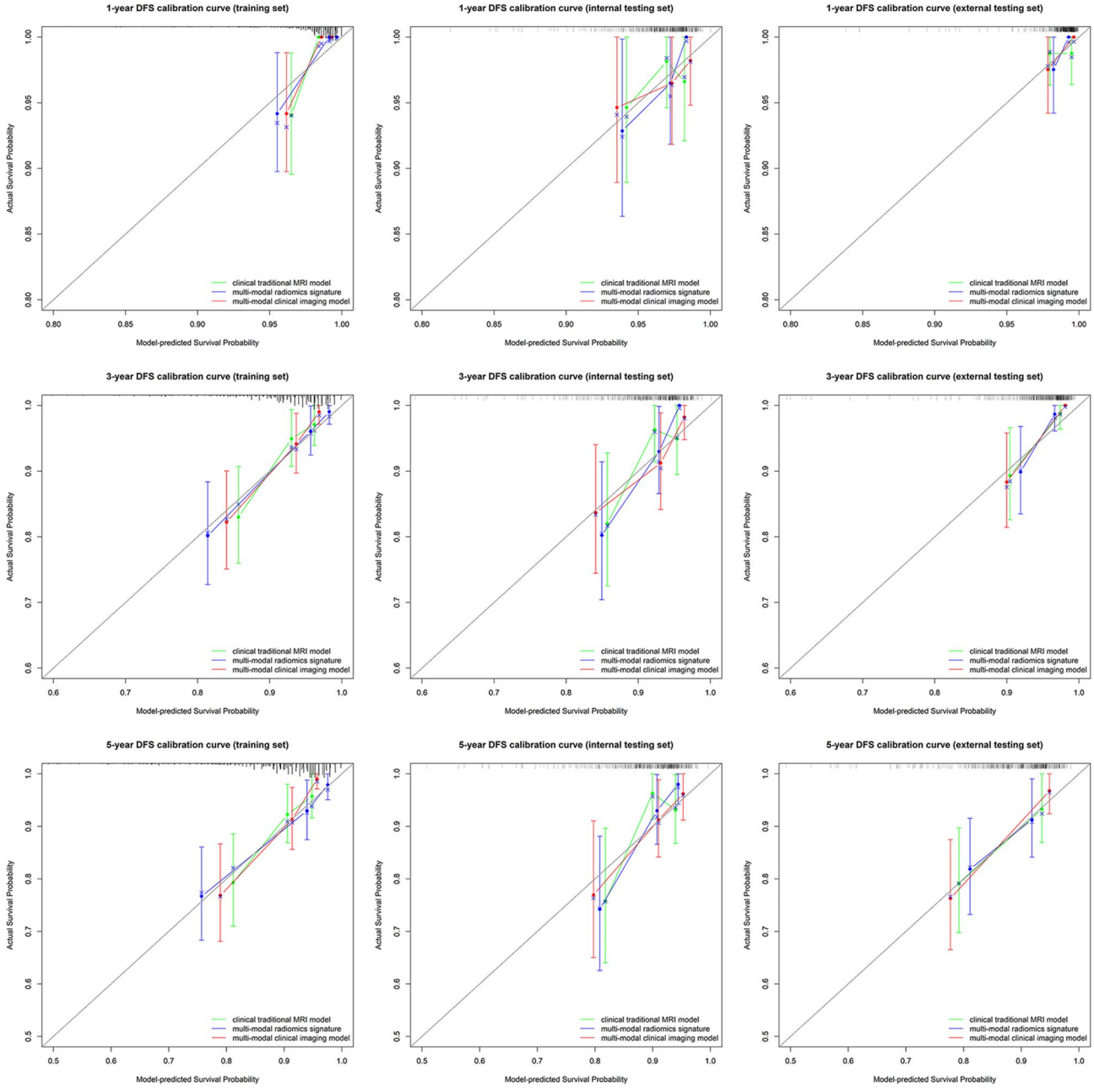

Fig. 5 Calibration curves show the calibration of the clinical traditional MRI model, the multi-modal radiomics signature, and the multi-modal clinical imaging model in terms of the agreement between the predicted and observed at 1-, 3-, and 5-year outcomes. The model-predicted and actual survival probability are plotted on the x- and y-axis, respectively. The diagonal gray line and colored line respectively represent a perfect estimation by an ideal model and the performance of models; the ideal model-predicted survival probability perfectly corresponds to the actual survival probability, while a closer alignment of the colored line with the diagonal gray line represents a better estimation of the model. MRI, magnetic resonance imaging; DFS, disease-free survival

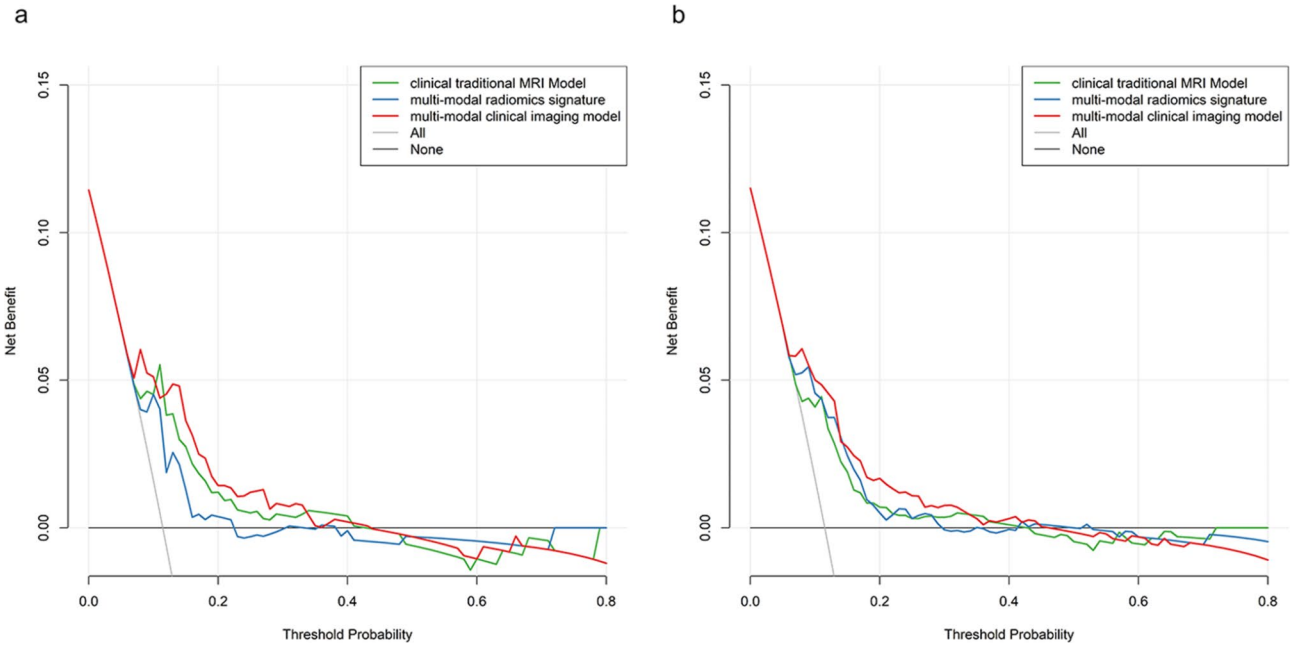

Fig. 6 Decision curve analysis in predicting 5-year disease-free survival in patients with invasive breast cancer for the clinical traditional MRI model, the multi-modal radiomics signature, and the multi-modal clinical imaging model when the internal and external testing sets were merged into a testing set (**a**), and when the training, internal and external testing sets were merged into a whole dataset (**b**). The net benefit, measured on the y-axis, was calculated by summing the benefits (true positive results) and subtracting the harms (false positive results), weighing the latter by a factor related to the relative harm of an undetected cancer compared with the harm of unnecessary treatment. The multi-modal clinical imaging model had the highest net benefit compared with both the other models and simple strategies such as follow-up of all patients (gray line) or no patients (horizontal black line) within the probability threshold range of 0.00-0.45 at which a patient would choose to undergo imaging follow-up. MRI, magnetic resonance imaging

evaluated in this study were incomplete. Third, the manually adjustments for the automatic segmentation-based tumor ROIs might introduce a segmentation bias during radiomics analysis, and the developed ASFs need further training and enhancement with new deep learning approaches and prospective multicenter data to improve their segmentation performance and reach higher levels of automation, thereby overcoming this problem. Thus, further investigations are needed to resolve these limitations.

Conclusions

In conclusion, the automatic segmentation-based multi-modal radiomics signature combining intra- and peri-radiomics signatures of US and MRI is a promising risk stratification biomarker for breast cancer, and the multi-modal clinical imaging model combining the multi-modal radiomics signature, the clinical and MRI features can improve individualized DFS prediction, which will likely assist in guiding decision-making related to patients diagnosed with breast cancer.

Abbreviations

3D	Three-dimensional
ASF	Automatic segmentation framework
AUC	Area under receiver operating characteristic curves
BIC	Bayesian information criterion
C-index	Harrell's concordance index
DCA	Decision curve analysis
DCE	Dynamic contrast enhanced
DFS	Disease-free survival
DSC	Dice similarity coefficient
ER	Estrogen receptor
HER2	Human epidermal growth factor receptor 2
HR	Hormone receptor
ICC	Intraclass correlation coefficient
LASSO	Least absolute shrinkage and selection operator
mASF	MRI ASF
MRI	Magnetic resonance imaging
PR	Progesterone receptor
Rad-score	Radiomics score
ROC	Receiver operating characteristic
ROI	Regions of interest
SYUCC	Sun Yat-sen University Cancer Center
TIC	Time-signal intensity curve
TNBC	Triple-negative breast cancer
uASF	US ASF
US	Ultrasound

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13058-024-01909-3.

Supplementary Material 1

Acknowledgements

We sincerely thank Han Jiao and Haiyan Lin of Electronics and Information, Sun Yat-sen University for their assistance in developing the automatic segmentation framework for breast cancer segmentation.

Author contributions
All authors contributed to the study conception and design. LX, XFT, XHJ, BYQ, BYC, and XFL did data collection. LX did the manual segmentation, HJ and HYL did the automatic segmentation and extracted radiomics features, XFT and XHJ helped with ROI delineation and confirmed ROIs. LX, XFT, XHJ, HLC, HJ and HYL did data analysis and interpretation. JHZ and LL provided study supervision. All authors drafted and revised the manuscript.

Funding
This study has received funding from Guangdong-Hongkong-Macao Applied Mathematics Center project of Guangdong Basic and Applied Basic Research Foundation (grant number 2021B1515310002).

Data availability
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This retrospective study was approved by the ethical review board (B2021-077-Y01) and the requirement for informed consent was waived.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Medical Imaging, First Affiliated Hospital of Gannan Medical University, Ganzhou, Jiangxi Province 341000, China
[2]Department of Ultrasound, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China
[3]Department of Medical Imaging, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China
[4]Department of Radiation Oncology, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China
[5]Department of Medical Imaging, Ganzhou People's Hospital, Ganzhou, Jiangxi Province 341000, China

Received: 16 April 2024 / Accepted: 21 October 2024
Published online: 12 November 2024

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. van Maaren MC, de Munck L, Strobbe LJA, Sonke GS, Westenend PJ, Smidt ML, et al. Ten-year recurrence rates for breast cancer subtypes in the Netherlands: a large population-based study. Int J Cancer. 2019;144(2):263–72.
3. Geurts YM, Witteveen A, Bretveld R, Poortmans PM, Sonke GS, Strobbe LJA, et al. Patterns and predictors of first and subsequent recurrence in women with early breast cancer. Breast Cancer Res Treat. 2017;165(3):709–20.
4. Blondeaux E, Poggio F, Del Mastro L. Role of dose-dense chemotherapy in high-risk early breast cancer. Curr Opin Oncol. 2019;31(6):480–5.
5. Johnston SRD, Harbeck N, Hegg R, Toi M, Martin M, Shao ZM, et al. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-, node-positive, high-risk, early breast cancer (monarchE). J Clin Oncology: Official J Am Soc Clin Oncol. 2020;38(34):3987–98.
6. Huang J, Tong Y, Chen X, Shen K. Prognostic factors and surgery for breast cancer patients with locoregional recurrence: an analysis of 5,202 consecutive patients. Front Oncol. 2021;11:763119.
7. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77.
8. Park H, Lim Y, Ko ES, Cho HH, Lee JE, Han BK, et al. Radiomics signature on magnetic resonance imaging: association with disease-free survival in patients with invasive breast cancer. Clin Cancer Res. 2018;24(19):4705–14.
9. Mazurowski MA, Saha A, Harowicz MR, Cain EH, Marks JR, Marcom PK. Association of distant recurrence-free survival with algorithmically extracted MRI characteristics in breast cancer. J Magn Reson Imaging. 2019;49(7):e231–40.
10. Chitalia RD, Rowland J, McDonald ES, Pantalone L, Cohen EA, Gastounioti A, et al. Imaging phenotypes of breast cancer heterogeneity in preoperative breast dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) scans predict 10-year recurrence. Clin Cancer Res. 2020;26(4):862–9.
11. Wang X, Xie T, Luo J, Zhou Z, Yu X, Guo X. Radiomics predicts the prognosis of patients with locally advanced breast cancer by reflecting the heterogeneity of tumor cells and the tumor microenvironment. Breast Cancer Res. 2022;24(1):20.
12. Yu F, Hang J, Deng J, Yang B, Wang J, Ye X, et al. Radiomics features on ultrasound imaging for the prediction of disease-free survival in triple negative breast cancer: a multi-institutional study. Br J Radiol. 2021;94(1126):20210188.
13. Wang H, Li X, Yuan Y, Tong Y, Zhu S, Huang R, et al. Association of machine learning ultrasound radiomics and disease outcome in triple negative breast cancer. Am J cancer Res. 2022;12(1):152–64.
14. Xiong L, Chen H, Tang X, Chen B, Jiang X, Liu L, et al. Ultrasound-based radiomics analysis for predicting disease-free survival of invasive breast cancer. Front Oncol. 2021;11:621993.
15. Guo Y, Wang Q, Guo Y, Zhang Y, Fu Y, Zhang H. Preoperative prediction of perineural invasion with multi-modality radiomics in rectal cancer. Sci Rep. 2021;11(1):9429.
16. Li ZY, Wang XD, Li M, Liu XJ, Ye Z, Song B, et al. Multi-modal radiomics model to predict treatment response to neoadjuvant chemotherapy for locally advanced rectal cancer. World J Gastroenterol. 2020;26(19):2388–402.
17. Zhao YF, Chen Z, Zhang Y, Zhou J, Chen JH, Lee KE, et al. Diagnosis of breast cancer using radiomics models built based on dynamic contrast enhanced mri combined with mammography. Front Oncol. 2021;11:774248.
18. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500–10.
19. Park JE, Park SY, Kim HJ, Kim HS. Reproducibility and generalizability in radiomics modeling: possible strategies in radiologic and statistical perspectives. Korean J Radiol. 2019;20(7):1124–37.
20. Rahimpour M, Saint Martin MJ, Frouin F, Akl P, Orlhac F, Koole M, et al. Visual ensemble selection of deep convolutional neural networks for 3D segmentation of breast tumors on dynamic contrast enhanced MRI. Eur Radiol. 2023;33(2):959–69.
21. Qiao M, Suo S, Cheng F, Hua J, Xue D, Guo Y, et al. Three-dimensional breast tumor segmentation on DCE-MRI with a multilabel attention-guided joint-phase-learning network. Comput Med Imaging Graphics: Official J Comput Med Imaging Soc. 2021;90:101909.
22. Ma M, Gan L, Jiang Y, Qin N, Li C, Zhang Y et al. Radiomics analysis based on automatic image segmentation of DCE-MRI for predicting triple-negative and nontriple-negative breast cancer. Computational and mathematical methods in medicine. 2021;2021:2140465.
23. Ma M, Gan L, Liu Y, Jiang Y, Xin L, Liu Y, et al. Radiomics features based on automatic segmented MRI images: prognostic biomarkers for triple-negative breast cancer treated with neoadjuvant chemotherapy. Eur J Radiol. 2022;146:110095.
24. Gan L, Ma M, Liu Y, Liu Q, Xin L, Cheng Y, et al. A clinical-radiomics model for predicting axillary pathologic complete response in breast cancer with axillary lymph node metastases. Front Oncol. 2021;11:786346.
25. Wang Z, Zhang H, Lin F, Zhang R, Ma H, Shi Y, et al. Intra- and peritumoral radiomics of contrast-enhanced mammography predicts axillary lymph node metastasis in patients with breast cancer: a multicenter study. Acad Radiol. 2023;30(Suppl 2):S133–42.
26. Sun Q, Lin X, Zhao Y, Li L, Yan K, Liang D, et al. Deep learning vs. radiomics for predicting axillary lymph node metastasis of breast cancer using ultrasound images: don't forget the peritumoral region. Front Oncol. 2020;10:53.
27. Jiang W, Meng R, Cheng Y, Wang H, Han T, Qu N, et al. Intra- and peritumoral based radiomics for assessment of lymphovascular invasion in invasive breast cancer. J Magn Reson Imaging. 2024;59(2):613–25.
28. Kwon BR, Shin SU, Kim SY, Choi Y, Cho N, Kim SM, et al. Microcalcifications and peritumoral edema predict survival outcome in luminal breast cancer treated with neoadjuvant chemotherapy. Radiology. 2022;304(2):310–9.
29. Mao N, Shi Y, Lian C, Wang Z, Zhang K, Xie H, 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):3207–19.
30. Jiang T, Song J, Wang X, Niu S, Zhao N, Dong Y, et al. Intratumoral and peritumoral analysis of mammography, tomosynthesis, and multiparametric MRI for predicting Ki-67 level in breast cancer: a radiomics-based study. Mol Imaging Biology. 2022;24(4):550–9.
31. Han X, Cao W, Wu L, Liang C. Radiomics assessment of the tumor immune microenvironment to predict outcomes in breast cancer. Front Immunol. 2021;12:773581.
32. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ. Strategies for subtypes–dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the primary therapy of early breast Cancer 2011. Annals Oncology: Official J Eur Soc Med Oncol. 2011;22(8):1736–47.
33. Whybra P, Zwanenburg A, Andrearczyk V, Schaer R, Apte AP, Ayotte A, et al. The image biomarker standardization initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights. Radiology. 2024;310(2):e231319.
34. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res. 2004;10(21):7252–9.
35. Gonen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. Biometrika. 2005;92:965–70.
36. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak. 2006;26(6):565–74.
37. Lee YW, Huang CS, Shih CC, Chang RF. Axillary lymph node metastasis status prediction of early-stage breast cancer using convolutional neural networks. Comput Biol Med. 2021;130:104206.
38. Zhang J, Saha A, Zhu Z, Mazurowski MA. Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radiogenomics. IEEE Trans Med Imaging. 2019;38(2):435–47.
39. Yu Y, Tan Y, Xie C, Hu Q, Ouyang J, Chen Y, et al. Development and validation of a preoperative magnetic resonance imaging radiomics-based signature to predict axillary lymph node metastasis and disease-free survival in patients with early-stage breast cancer. JAMA Netw open. 2020;3(12):e2028086.

Publisher's note