Clinical Cancer Research

Radiomics of Multiparametric MRI for Pretreatment Prediction of Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer: A Multicenter Study

Zhenyu Liu[1,2], Zhuolin Li[3], Jinrong Qu[4], Renzhi Zhang[5], Xuezhi Zhou[1,6], Longfei Li[1,7], Kai Sun[1,6], Zhenchao Tang[1], Hui Jiang[4], Hailiang Li[4], Qianqian Xiong[8], Yingying Ding[3], Xinming Zhao[5], Kun Wang[8], Zaiyi Liu[9], and Jie Tian[1,2,6,10]

Abstract

Purpose: We evaluated the performance of the newly proposed radiomics of multiparametric MRI (RMM), developed and validated based on a multicenter dataset adopting a radiomic strategy, for pretreatment prediction of pathologic complete response (pCR) to neoadjuvant chemotherapy (NAC) in breast cancer.

Experimental Design: A total of 586 potentially eligible patients were retrospectively enrolled from four hospitals (primary cohort and external validation cohort 1–3). Quantitative imaging features were extracted from T2-weighted imaging, diffusion-weighted imaging, and contrast-enhanced T1-weighted imaging before NAC for each patient. With features selected using a coarse to fine feature selection strategy, four radiomic signatures were constructed based on each of the three MRI sequences and their combination. RMM was developed based on the best radiomic signature incorporating with independent clinicopathologic risk factors. The performance of RMM was assessed with respect to its discrimination and clinical usefulness, and compared with that of clinical information–based prediction model.

Results: Radiomic signature combining multiparametric MRI achieved an AUC of 0.79 (the highest among the four radiomic signatures). The signature further achieved good performances in hormone receptor–positive and HER2-negative group and triple-negative group. RMM yielded an AUC of 0.86, which was significantly higher than that of clinical model in two of the three external validation cohorts.

Conclusions: The study suggested a possibility that RMM provided a potential tool to develop a model for predicting pCR to NAC in breast cancer.

[1]CAS Key Laboratory of Molecular Imaging, Institute of Automation, Beijing, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Department of Radiology, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital), Kunming, Yunnan, China. [4]Department of Radiology, Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou, Henan, China. [5]Department of Diagnostic Radiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. [6]Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, China. [7]Collaborative Innovation Center for Internet Healthcare, Zhengzhou University, Zhengzhou, Henan, China. [8]Department of Breast Cancer, Cancer Center, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, Guangzhou, China. [9]Department of Radiology, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, Guangzhou, China. [10]Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, China.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Z. Liu, Z. Li, J. Qu, and R. Zhang contributed equally to this article.

Corresponding Authors: Jie Tian, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. Phone: 86-10-82618465; Fax: 86-10-62527995; E-mail: tian@ieee.org; Zaiyi Liu, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, Guangzhou 510080, China. E-mail: zyliu@163.com; and Kun Wang, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, Guangzhou 510080, China. E-mail: gzwangkun@126.com

Clin Cancer Res 2019;25:3538–47

doi: 10.1158/1078-0432.CCR-18-3190

Introduction

Breast cancer has the highest incidence among cancers in women worldwide (1). Neoadjuvant chemotherapy (NAC) has been established as a standard treatment of care for most breast cancers, especially locally advanced breast cancer (2). NAC is able to downstage cancer, reduce metastasis, detect drug sensitivity, and improve the possibility of breast-conserving therapy (3, 4). Ideally, it could imply an extremely favorable disease-free and overall survival when a pathologic complete response (pCR) is achieved after NAC (5). Thereafter, pCR could be proposed as a surrogate early clinical endpoint for long-term survival (6); however, there is still no standard method to predict responses to NAC. As the outcome of NAC tends to be varied across histopathologic and molecular characteristics (7), it makes the quantitative pretreatment prediction of pCR for better treatment planning even more challenging.

Various prediction methods have been proposed to predict the responses to NAC in patients with breast cancer, including physical examination and medical imaging tests like mammography, ultrasonography, diffuse optical spectroscopic (8), breast MRI, and PET/CT (9, 10). Although MRI-based methods cannot detect pCR with adequate accuracy (11), it is currently the most accurate method for determining response to NAC (9, 10, 12). Specifically, contrast-enhanced MRI was considered as the most reliable technique for evaluating the responses to NAC (13, 14) at present, as its measurements of perfusion and permeability of

Translational Relevance

In this study, we developed and validated radiomics of multiparametric MRI (hereafter RMM) based on a multicenter dataset for pretreatment prediction of pathologic complete response (pCR) to neoadjuvant chemotherapy (NAC) in breast cancer. Radiomic signature combining T2-weighted imaging, diffusion-weighted imaging, and contrast-enhanced T1-weighted imaging MRI showed good performance within the primary and external validation cohorts. Moreover, the radiomic signature yielded relatively well performances in hormone receptor-positive and HER2-negative group and triple-negative group. Furthermore, RMM incorporating radiomic signature and clinical information showed improved performance in predicting pCR to NAC compared with a prediction model based on clinical information in the multicenter dataset. The study suggested a possibility that RMM provided a potential tool to develop a model for predicting pCR to NAC in breast cancer.

tissue microvessels are sensitive to angiogenic changes. However, it is still difficult to use contrast-enhanced MRI for pretreatment predicting of pCR. Diffusion-weighted imaging (DWI) quantitatively measured apparent diffusion coefficients (ADC), which reflect the diffusivity of water and provided information on the integrity of cell membranes and tumor cellularity. It is sensitive to chemotherapy induced intratumoral changes, Hence, DWI may provide complementary information for predicting responses to chemotherapy (13, 15). Thus, multiparametric MRI combining conventional T2-weighted imaging (T2WI), DWI, and contrast-enhanced T1-weighted imaging (T1+C) may achieve better and a more robust performance in predicting responses to NAC.

Radiomics is a rapidly emerging field involving the extraction of numerous quantitative features from multimodality medical images to determine relationships between the features and the underlying pathophysiology (16–19). Based on the concept that biomedical images contain information that may reflect underlying pathophysiology and their relationships could be revealed via quantitative image analyses (17), radiomics turns medical images into minable data to improve diagnostic (20, 21), prognostic (22), and predictive (23) accuracy, bridging the gap between medical imaging and personalized medicine (24). In addition, radiomics can combine clinical information and histopathologic and molecular characteristics with multiple imaging features to deliver more accurate medical care (25). Radiomics can also be used for assessing responses to antitumor therapy. Recent studies have proposed radiomic approaches for predicting pCR to neoadjuvant therapy in rectal cancer (26) and assessing response to immunotherapy in solid tumors (27). Moreover, radiomics was utilized in two studies with relatively small cohorts for predicting pCR to NAC in breast cancer (28, 29). These studies demonstrated the feasibility and potential benefits of using radiomics in pCR prediction in breast cancer. However, these studies have small sample sizes and are not validated based on multicenter dataset, with lack of comparison with results of clinical information-based methods, and these may limit their clinical application.

In this study, we proposed to develop a radiomic model based on multiparametric MRI and clinical information for pretreatment pCR prediction in breast cancer, named radiomics of multiparametric MRI (RMM), and validated it with a multicenter dataset. We hypothesized that RMM was with the potential in the prediction of pCR to NAC in patients with breast cancer.

Materials and Methods

Study design

This was a multicenter study with patients retrospectively enrolled from four Chinese hospitals in different regions of China. A new approach named RMM integrating pretreatment multiparametric MRI (T2WI, DWI, and T1+C) and clinical information was proposed to predict pCR to NAC in patients with primary invasive breast cancer. Histopathologic examination of surgically resected specimens was used as the reference standard, and RMM was compared with a clinical information-based prediction model and MRI prediction models constructed with T2WI, DWI, and T1+C. This multicenter study was conducted in accordance with the Declaration of Helsinki and was approved by the ethics committee of each participating hospital, with the requirement for informed consent waived.

Patients

The inclusion criteria were as follows: (i) the patient had biopsy-proven unilateral primary invasive breast cancer without distant metastasis; (ii) the patient received complete NAC with no prior treatment; (iii) surgery was performed after the completion of NAC, after which pCR was confirmed by postoperative pathologic examination; and (iv) pretreatment breast MRI was conducted before biopsy, including T2WI, DWI, and T1+C. The exclusion criteria were (i) the patient was undergoing biopsy at an external institution and pretreatment pathologic results were not available; (ii) the patient did not complete NAC or had nonstandard treatment (mainly referring to HER2-positive tumors that were not treated with trastuzumab); (iii) the patient was undergoing surgery at an external institution, or pCR was not assessed; (iv) lack of pretreatment T2WI, DWI, or T1+C; (v) insufficient MRI quality to obtain measurements (e.g., owing to motion artifacts); and (vi) the patient had unilateral multifocal cancers, and the correlation between the tumor in MRI and postoperative pathologic examination was uncertain.

The data set with the most enrolled patients was used as the primary cohort (PC) to reduce any form of overfitting or bias in the analysis, and the other three data sets were used as independent validation cohorts (VC1–VC3).

MRI data acquisition and IHC

Breast MR examination for each patient was performed before biopsy and within 1 to 2 weeks before NAC. Fat-suppressed T2WI, DWI, and T1+C with fat suppression were acquired for each patient in the four cohorts. An axial fat-suppressed T2WI sequence was acquired before contrast medium administration. Then, an initial fat-saturated T1WI precontrast scan was collected before T1+C images scanning, and T1+C images were then acquired following the intravenous injection of gadolinium contrast agent. Finally, axial DWI images were obtained using two *b* values (0 and 1,000 s/mm^2). The detailed parameters of MR images acquisition of the four hospitals can be found in the Supplementary Information and Supplementary Table S1.

The status of estrogen receptor (ER), progesterone receptor (PR), and HER2, and the Ki67 index were determined by IHC.

We defined tumors with <1% of tumor cells with nuclear staining as ER/PR negative and ≥1% of tumor cells with nuclear staining as ER/PR positive (30); the cutoff for Ki67 was set at 20%. For HER2, tumors with IHC staining of 0 or 1+ was defined as HER2 negative whereas tumors with IHC staining of 3+ was defined as HER2 positive. For tumors with IHC staining 2+, further confirmation was obtained with molecular testing (ISH testing): ISH non-amplified results were defined as HER2 negative and ISH amplified results were considered HER2 positive.

NAC and pathologic assessment of response

All patients received four cycles, six cycles, or eight cycles of NAC prior to breast surgery. Although there were some differences among the different hospitals, the treatment protocol and timeline followed the National Comprehensive Cancer Network (NCCN) guideline (2). The NAC regimens were either taxane-based, anthracycline-based, or anthracycline and taxane-based (detailed NAC regimens in each cohort was shown in Supplementary Table S2). Additionally, HER2-positive patients also received trastuzumab (8 mg/kg as the loading dose and 6 mg/kg as the maintenance dose).

Standard histopathologic analysis was conducted in each participating hospital for the pathologic assessment of response to NAC. Surgically resected specimens were fixed in 10% neutral buffered formalin and processed overnight in standard tissue processors and slides were cut at 5 μm and stained an automated staining system. The histopathologically examination and analysis were dedicated by breast pathologists (with at least a 10-year experience in breast pathology) who were blinded to the MRI data from the participating hospitals. pCR was defined as the absence of residual invasive carcinoma in the specimen (residual ductal carcinoma *in situ* could be present) and the absence of lymph node invasion in the ipsilateral sentinel node or lymph nodes removed during axillary dissection (ypT0/isN0; refs. 31–33).

Tumor masking and inter- and intraobserver reproducibility evaluation

Pretreatment MRI data from all participating hospitals were collated for tumor masking and features extraction. The regions of interest (ROI) were delineated manually via the itk-SNAP software (www.itksnap.org) on each slice of the T2WI, DWI (*b*-value of 1,000 s/mm^2), and T1+C (the peak enhanced phase of the multiphase contrast enhanced MRI selected according to the time intensity curve) data by excluding the necrosis, air, and calcification area. Because of the higher resolution of DWI in comparison to ADC maps, ROIs were detected with a *b*-value of 1,000 s/mm^2 first, and then copied to the corresponding ADC maps for further analysis. ROIs of breast cancer were manually drawn along the contour of the tumor on T2WI (slightly high signal) and T1+C (enhanced region) containing the surrounding chords and burrs, and ROIs were also placed on the high signal intensity region on DWI (*b*-value of 1,000 s/mm^2).

Four radiologists (1 from each participating hospital) with at least 10 years' experience in breast MR imaging were chiefly responsible for the evaluation of tumor masking. Inter- and intra-observer reproducibility of tumor masking and radiomic feature extraction were initially analyzed with the T2WI data of 30 randomly selected patients for ROI-based radiomic feature generation in a blinded fashion by these 4 radiologists.

To ensure reproducibility, each radiologist repeated the tumor masking and generation of radiomic features twice with an interval of at least 1 month, following the same procedure. Intra-class correlation coefficients (ICCs) were utilized for evaluating the intra- and inter-observer agreement in terms of feature extraction. We interpreted an ICC of 0.81-1.00 as almost perfect agreement, 0.61-0.80 as substantial agreement, 0.41-0.60 as moderate agreement, 0.21-0.40 as fair agreement, and 0-0.20 as poor or no agreement (34). An ICC greater than 0.6 was considered a mark of satisfactory inter- and intraobserver reproducibility.

To ensure the accuracy of tumor masking, the tumor masks were evaluated by other radiologists from the same hospital for each hospital, following the same guideline describing how to define the boundary of tumors.

Radiomic feature extraction

Radiomic feature extraction was performed with MATLAB 2016b (Mathworks) using a toolbox developed in-house toolbox. Each MRI scan of each patient was normalized with *Z*-scores in order to get a standard normal distribution of image intensities. Then, four groups of imaging features were extracted from each normalized pretreatment MRI scan with manually segmented ROIs: Group 1 comprised eight shape- and size-based features, Group 2 comprised 17 first-order statistical features, Group 3 was comprised of 90 textural features and 4,535 wavelet features (4,280 features of Gabor-bank wavelet filtered images and 155 features of Law's filtered images). Group 1 quantitatively described the three-dimensional size and shape of the tumor. Group 2 consisted of quantified tumor intensity characteristics with first-order statistics calculated from the histogram of all tumor intensities. Group 3 comprised textual features based on the quantification of intratumoral heterogeneity (i.e., differences in texture observed within the tumor volume). Group 4 incorporated the calculated textural features from the wavelet decompositions of the original images, thereby focusing on the various wave lengths and different feature orientation within the tumor volume. All of these features have generally been used in previous radiomic studies (20, 22, 26). The final feature set comprised 4,650 features for each MR sequence (T2WI, ADC, and T1+C), resulting in a total of 13,950 radiomic features per patient. Details of all feature-extracting algorithms are provided in the Supplementary Information.

Radiomic signature construction and validation

Radiomic signatures based on single MR sequences and multiparametric MRI were constructed within the PC. Thus, four radiomic signatures were generated (based on T2WI, DWI, T1+C, and multiparametric MRI combining the above three sequences), and all of them were validated with the validation cohorts.

For the construction of radiomic signatures, the same coarse to fine feature selection strategy was utilized to reduce any bias of results and potential overfitting. First, univariate analysis was performed with the Mann–Whitney *U* test to compare radiomic features between pCR and non-pCR patients. All features were ranked according to the *P*-value from the Mann–Whitney *U* test in ascending order, and the top 5% of the features were used for further analysis. Second, the Pearson correlation coefficient between each pair of features was computed (denoted as r thereafter). All pairs of features with $|r| > 0.85$ were detected, and the feature in each of these pairs with the larger *P*-value from the Mann–Whitney *U* test was deleted from the feature set. Finally, a random forest based feature selection method named Boruta (35)

was used to detect the key features for pCR prediction. Boruta could select all relevant features for prediction instead of only the nonredundant ones.

With the selected key features, supporting vector machine (SVM) models were used to construct radiomic signatures for pCR prediction. SVM models with radial basis function kernel were trained based on the PC, and a 10-fold cross validation was used to determine the parameters of the SVM models. After a model was trained, the radiomic score for each patient in the validation cohort was computed. To assess the quantitative prediction performance of the four radiomic signatures in the primary and validation cohorts, ROC curves and area under the ROC curves (AUC) were calculated in all these four cohorts.

Performance of radiomic signature according to breast cancer subtype

Patients in each cohort were divided into subgroups according to their breast cancer subtype. Patients with ER and/or PR positive and HER2-negative breast cancer were grouped into the hormone receptor (HR)-positive, HER2 negative (HR+ and HER2−) subgroup. Patients with HER2-positive breast cancer were grouped into the HER2+ subgroup and the remaining patients with ER−, PR−, and HER2− breast cancer were grouped into the triple-negative (TN) subgroup. The radiomic signature for each breast cancer subtype was obtained by training the model with eight selected features from the multiparametric MRI based on patients of the corresponding subtype in the PC. The prediction performance was tested with patients of the corresponding subtype in the three validation cohorts.

RMM and its overall performance

The radiomic signature of multiparametric MRI was applied with age, stage, ER status, PR status, HER2 status, and Ki-67 status using multivariable logistic regression analysis to develop RMM for pCR prediction based on the PC. Backward step-wise selection was applied by using the likelihood ratio test with Akaike's information criterion as the stopping rule to select correlated factors of pCR (36, 37). The prediction results of RMM in the validation cohorts were used to evaluate the performance of RMM.

A clinical model was also constructed using multivariable logistic regression analysis with age, stage, ER status, PR status, HER2 status, and Ki-67 status. Thus, the prediction performance of RMM was compared with that of the clinical model and radiomic signatures.

Statistical analysis

Descriptive statistics were summarized as mean ± SD. Comparisons between groups were made with the t test or Mann–Whitney U test, when appropriate, for quantitative variables and with the χ^2 test or Fisher's test for qualitative variables. AUC and 95% confidence interval (CI) calculated with Statistical Product and Service Solutions (SPSS) were used to evaluate the performance of each model for pCR prediction. Differences between various AUCs were compared with the DeLong test (38). All statistical tests were two sided and P values less than 0.05 indicating statistical significance. The statistical analyses were performed using SPSS software, version 21 (SPSS).

Results

Clinical characteristics

As shown in Figure 1, a total of 586 potentially eligible patients were consecutively enrolled in this study from the four participating hospitals, and 172 patients were excluded according to the exclusion criteria. Thus, 414 patients were finally enrolled for further analysis. The dataset from Guangdong General Hospital had the highest number of eligible patients (128) and was used as the PC. The clinical characteristics of all patients are summarized in Table 1 (Supplementary Table S3).

The pCR rate in the four cohorts was between 15% and 43%. In all four cohorts, no significant difference was detected between the pCR and non-pCR groups in terms of age, stage and Ki-67 status ($P > 0.05$) was detected. Meanwhile, pCR was found to be

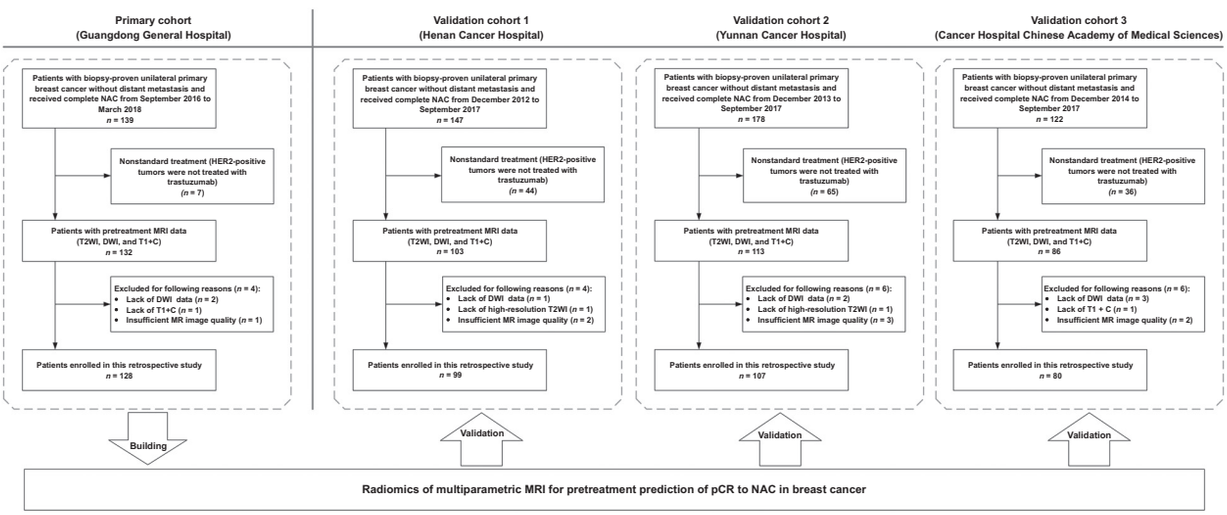

Figure 1.

Patient recruitment and study design. In total, 414 of 586 patients with pretreatment multiparametric MRI from four Chinese hospitals were enrolled in this study for model construction and validation.

Table 1. Demographic comparison between pCR and non-pCR groups in the primary and validation cohorts

	PC ($N = 128$)			Validation cohort 1 ($N = 99$)			Validation cohort 2 ($N = 107$)			Validation cohort 3 ($N = 80$)		
Characteristics	**pCR ($n = 56$)**	**Non-pCR ($n = 72$)**	***P***	**pCR ($n = 16$)**	**Non-pCR ($n = 83$)**	***P***	**pCR ($n = 28$)**	**Non-pCR ($n = 79$)**	***P***	**pCR ($n = 12$)**	**Non-pCR ($n = 68$)**	***P***
Age (years, mean ± SD)	47.52 ± 9.53	48.16 ± 8.71	0.473	50.13 ± 10.01	45.28 ± 7.97	0.283	45.86 ± 7.42	46.03 ± 9.26	0.284	44.32 ± 11.60	44.32 ± 11.60	0.440
Stage (%)			0.560			0.822			1.000			0.212
I	6 (10.71)	7 (9.73)		0 (0.00)	1 (1.20)		0 (0.00)	0 (0.00)		1 (8.33)	1 (1.47)	
II	43 (76.79)	51 (70.83)		8 (50.00)	37 (44.58)		21 (75.00)	57 (72.15)		2 (16.67)	23 (33.82)	
III	7 (12.50)	14 (19.44)		8 (50.00)	45 (54.22)		7 (25.00)	22 (27.85)		9 (75.00)	44 (64.71)	
ER status (%)			0.001*			0.001*			0.179			0.001*
Positive	25 (44.64)	56 (77.78)		6 (37.50)	66 (79.52)		19 (67.86)	65 (82.28)		1 (8.33)	41 (60.29)	
Negative	31 (55.36)	16 (22.22)		10 (62.50)	17 (20.48)		9 (32.14)	14 (17.72)		11 (91.67)	27 (39.71)	
PR status (%)			0.001*			0.002*			0.204			0.532
Positive	18 (32.14)	48 (66.67)		5 (31.25)	61 (73.49)		18 (64.29)	62 (78.48)		5 (41.67)	38 (55.88)	
Negative	38 (67.86)	24 (33.33)		11 (68.75)	22 (26.51)		10 (35.71)	17 (21.52)		7 (58.33)	30 (44.12)	
HER2 status (%)			0.001*			0.099			0.449			0.031*
Positive	36 (64.29)	21 (29.17)		6 (37.50)	15 (18.07)		4 (14.29)	6 (7.59)		5 (41.67)	9 (13.24)	
Negative	20 (35.71)	51 (70.83)		10 (62.50)	68 (81.93)		24 (85.71)	73 (92.41)		7 (58.33)	59 (86.76)	
Ki-67 status (%)			0.365			0.348			0.372			0.201
Positive	48 (85.71)	57 (79.17)		16 (100.00)	74 (89.16)		19 (67.86)	44 (55.70)		12 (100.00)	57 (83.82)	
Negative	8 (14.29)	15 (20.83)		0 (0.00)	9 (10.84)		9 (32.14)	35 (44.30)		0 (0.00)	11 (16.18)	
Cancer subtype (%)			0.001*			0.001*			0.064			0.014*
HR+ and HER2−	9 (16.07)	42 (58.33)		3 (18.75)	59 (71.08)		17 (60.71)	65 (82.28)		2 (16.66)	39 (57.35)	
HER2+	36 (64.29)	21 (29.17)		6 (37.50)	15 (18.07)		4 (14.29)	6 (7.59)		5 (41.67)	9 (13.24)	
Triple-negative	11 (19.64)	9 (12.50)		7 (43.75)	9 (10.85)		7 (25.00)	8 (10.13)		5 (41.67)	20 (29.41)	

NOTE: χ^2 or Fisher's exact tests, as appropriate, were used to compare the differences in categorical variables, whereas a two-sample *t* test was used to compare the differences in age. The ER and PR threshold value for level was ≤1%, and the threshold value for Ki-67 was ≤20%.

*$P < 0.05$.

significantly associated with HER2 status only in the PC and VC three but not in the other two validation cohorts. ER status and PR status showed significant differences between the two groups in the PC and in two and one of the validation cohorts, respectively. The results suggested that pCR could be correlated with the status of molecular receptors, which was consistent with that of previous studies (3, 5). Thereafter, a clinical model including the information for pCR prediction was constructed subsequently as a baseline to evaluate the RMM proposed in this study.

Feature discovery and radiomic signature construction and validation

Satisfactory inter- and intraobserver reproducibility of tumor masking and radiomic feature extraction was achieved with ICC > 0.6 both among the masks from the four radiologists at baseline and between the masks from the same radiologist at baseline and at least 1 month later.

With the coarse to fine feature selection strategy, 7, 8, and 3 imaging features were finally selected from the T2WI, ADC maps, and T1+C respectively for the construction of radiomic signatures based on a single sequence, and eight imaging features were selected from the full feature set including features of the above three sequences to construct the radiomic signature based on multiparametric MRI (Supplementary Table S4). In the final step with Boruta, a set of features that have confirmed or tentative association with pCR was selected from each of the four feature sets for signatures construction based on the PC (Fig. 2).

An SVM model was then constructed using the selected features for each sequence and multiparametric MRI as a radiomic signature. The ROC curves and AUCs of the four radiomic signatures in the PC and the external validation cohorts are shown in Figure 3A–D) and Supplementary Table S5. The multiparametric MRI-based radiomic signature showed an improved performance, with an AUC of 0.79 over the single sequence radiomic signatures based on T2WI (AUC = 0.69, $P = 0.042$), ADC (AUC = 0.69, $P = 0.053$), and T1+C (AUC = 0.64, $P = 0.002$) in the PC. Although the performances of all radiomic signatures dropped in the validation cohorts, the multiparametric MRI-based radiomic signature outperformed all the single sequence radiomic signatures with AUCs larger than or close to 0.7.

Performance of radiomic signature according to breast cancer subtype

The ROC curves of radiomic signature based on multiparametric MRI within the three breast cancer subtypes in the primary and validation cohorts are shown in Figure 4 (Supplementary Table S6). Prediction within the HR+ and HER2−, and the TN subgroups achieved good performance in the primary and validation cohorts. In the HER2+ subgroup, although the signature had AUCs of 0.7 in the PC and 0.79 in VC1, the AUCs in VC2 and VC3 decreased to 0.58 and 0.62, respectively. Of particular note, the AUCs of the radiomic signature for the TN subgroup in the PC reached 0.96 and larger than or closed to 0.8 in all the three validation cohorts.

RMM and its overall performance

The radiomic signature of multiparametric MRI, PR status, and HER2 status were identified as independent predictors for the pCR prediction model RMM (Fig. 3E–H; Supplementary Table S7) with multivariable logistic regression analysis. As shown in Figure

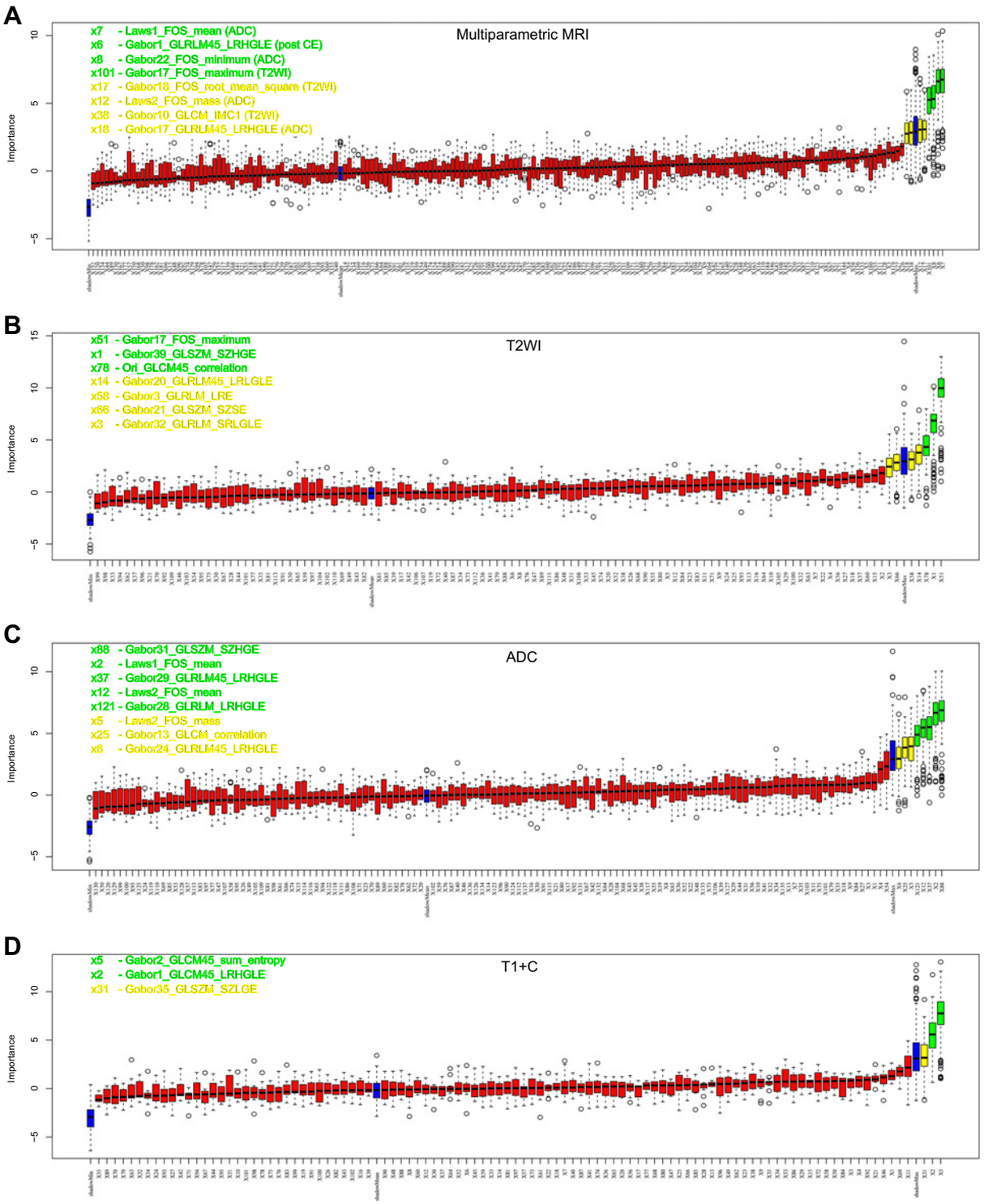

Figure 2.

Feature selection for construction of radiomic signatures with Boruta algorithm. Blue boxplots depict minimal, average, and maximum importance of a shadow attribute. Yellow boxplots correspond to tentative features, whereas green ones represent confirmed features. Tentative and confirmed features were both selected for further analysis. **A,** Eight imaging features were selected from the combination of T2WI, ADC, and T1+C for the radiomic signature. **B,** Seven imaging features were selected from the sequence of T2WI for the radiomic signature. **C,** Eight imaging features were selected from the ADC maps for the radiomic signature. **D,** Three imaging features were selected from the sequence of T1+C for the radiomic signature.

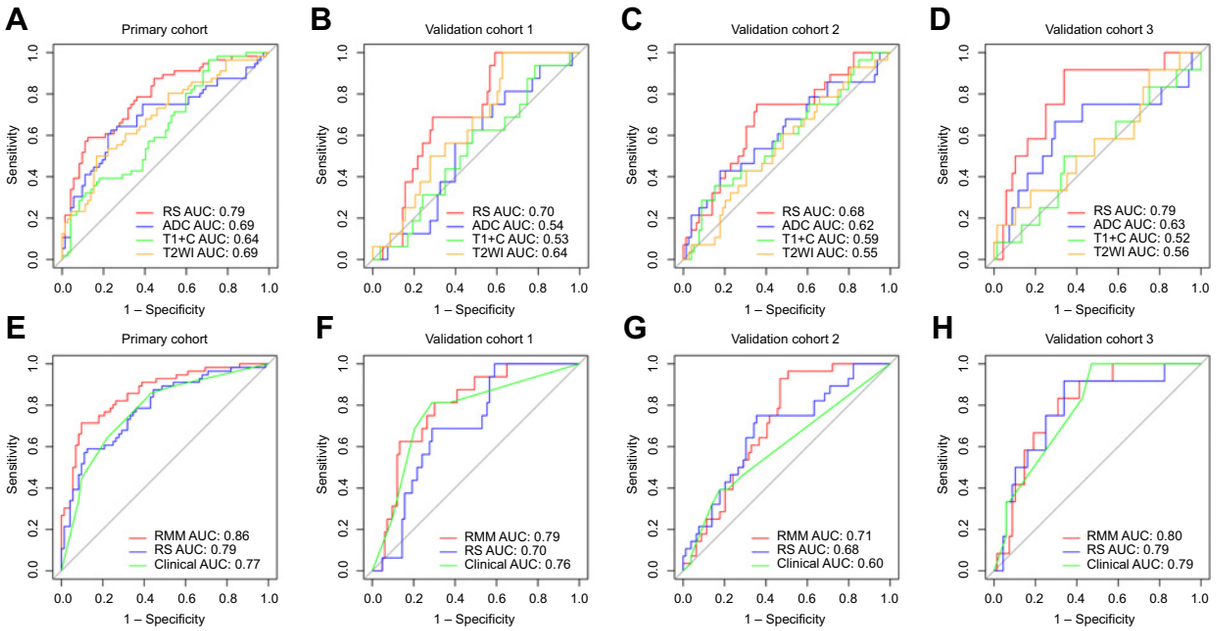

Figure 3.

ROC curves among different radiomic signatures and among different models. **A–D,** ROC curves of radiomic signatures based on single sequence and multiparametric MRI for pCR prediction in the primary and validation cohorts. **E–H,** ROC curves of RMM, radiomic signatures, and clinical model for pCR prediction in the primary and validation cohorts.

3E, RMM yielded a good prediction performance with an AUC of 0.86 in the PC. With a similar backward step-wise selection method, ER status, PR status, and HER2 status were identified as predictors for the clinical model of pCR prediction, which achieved a performance with an AUC of 0.77 in the PC (Fig. 3E). Using the DeLong test, the RMM showed significantly higher AUC than that of both the multiparametric MRI-based radiomic signature and the clinical model ($P < 0.05$) in the PC. Good performance was also observed for pCR prediction in validation cohorts; although the AUC of RMM dropped marginally, the AUCs were larger than 0.7 in all three external validation cohorts, which were the highest of the three models, although the performances of clinical model and multiparametric MRI based radiomic signature dropped more than RMM. The AUCs comparison in the validation cohorts showed that RMM achieved significantly higher AUC than clinical model in VC1 and VC2, and it outperformed multiparametric MRI-based radiomic signature in VC1 (Supplementary Table S8).

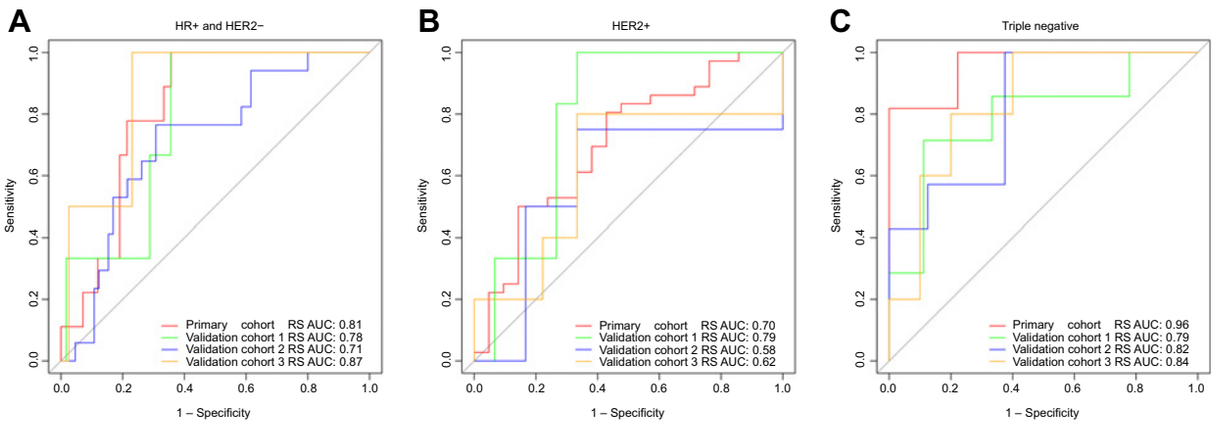

Figure 4.

ROC curves of radiomic signature based on multiparametric MRI according to breast cancer subtypes in the primary and validation cohorts. **A,** ROC curve of radiomic signature based on multiparametric MRI for pCR prediction in the subtype of HR+ and HER2−. **B,** ROC curve of radiomic signature based on multiparametric MRI for pCR prediction in the subtype of HER2+. **C,** ROC curve of radiomic signature based on multiparametric MRI for pCR prediction in the subtype of triple negative.

Discussion

In this multicenter study, we investigated the ability of pretreatment multiparametric MRI-based radiomic analysis to predict pCR to NAC in patients with breast cancer. A model named RMM was proposed with better performance in the primary and majority of external validation cohorts compared with prediction model based on clinical information. The outperformance of RMM indicated that combining clinical information with multiparametric MRI could be helpful for the pretreatment prediction of pCR to NAC in breast cancer.

The prediction ability of RMM was significantly improved compared with that of other models including clinical model and radiomic signatures based on multiple and single sequences ($P < 0.05$) in the PC and two of the three external validation cohorts. Although the AUCs of the clinical model and multiparametric MRI-based radiomic signature reached 0.77 and 0.79 in the PC, they were unable to perform robustly in all validation cohorts. Specifically, the AUC of the clinical model in VC2 was 0.60, which was significantly lower than that of RMM ($P < 0.05$). The distribution of the receptor status of ER, PR, and HER2 in the patients in VC2, which were incorporated in the clinical model, was significantly different from that of the PC. This may be the reason why the clinical model had a poor performance in this cohort. Previous studies have pointed out that different molecular subtypes could achieve different rates of pCR (5, 7), indicating that a different distribution of cancer subtypes could mean different treatment effects and different rates of pCR, hence the clinical model did not work well on a dataset with a different patient distribution. As the clinical information may only take into consideration certain aspects of the tumor, multiparametric MRI may better reflect all information on the tumor (17). Thus, when we combined the clinical information and imaging features for RMM, better performance was achieved. The high-dimensional imaging features may be able to acquire more detailed information about the tumor that cannot be detected easily by the naked eye, including the molecular subtype of breast cancer, and can more comprehensively describe the tumor (18, 24).

The eight radiomic features selected for the final multiparametric MRI radiomic signature and RMM comprised one feature from T1+C, three features from T2WI, and four features from ADC (computed with DWI). All eight features were obtained from Gabor-bank or Law's filtered images, which are high-dimensional features that cannot be easily deciphered by humans but hold more detailed information about cancer and are more sensitive for treatment evaluation. The results suggested that combining multiple MRI sequences allowed the detection of more detailed information on the tumor. Although T2WI provided morphological features of breast cancer, limited information about the response to NAC could be detected. Recent studies have suggested that contrast enhanced MRI could be the best tool to predict pCR to NAC so far (28, 29, 39), and DWI was considered a potential tool for measuring treatment response in breast cancer (15, 40). The combination of T2WI, DWI, and T1+C allows the detection of morphological information, water diffusion properties in tissue (also cellularity and interstitial water mobility from ADC maps), and permeability of tissue microvessels at the same time. As multiparametric MRI was successfully applied in treatment evaluation in patients with rectal cancer recently (26, 41), it is natural to investigate the performance of multiparametric MRI in predicting pCR in patients with breast cancer. As expected, RMM achieved better performance in this multicenter study.

We also found that multiparametric MRI-based radiomic signature can make good predictions of pCR to NAC for patients in the HR+ and HER2− subgroup and especially for patients in the TN subgroup with high AUCs in not only the PC but also the three external validation cohorts. The prediction performance was similar to a previous study on patients in the HR+ and HER2− subgroup, but it did not have external validation (28). Our results provided additional evidence supporting the use of MRI in predicting pCR. Although patients with TN breast cancers usually achieved pCR more frequently, their response to NAC could be more accurately assessed with posttreatment MRI (42). We further demonstrated that pretreatment multiparametric MRI-based radiomics could more accurately predict pCR in patients in the TN subgroup. For the patients in the HER2+ subgroup, the radiomic signature performed well in the PC and VC1 but had decreased AUCs in VC2 and VC3. As some patients who received nonstandard treatment (not treated with trastuzumab) were excluded from the study, only a small sample of patients with this subtype were included, which could impact the rate of pCR and prediction in this subgroup, and also the performance of radiomic signature decreased in the smaller cohorts of VC2 and VC3 in this subtype. This could be a limitation of the present study. With more patients receiving standard treatment, a better prediction model could be obtained.

There were still some limitations in the study, mainly due to the limited population size, and the unbalanced patient distribution. Although the proportion of patients who reached pCR in all four cohorts was in the normal range (7), it was significantly higher in the dataset from Guangdong General Hospital. This may be due to the administration of standard treatment to patients with HER2+ breast cancer (trastuzumab in addition to the routine NAC), as the proportion of patients who reached pCR could be much higher in patients with HER2+ breast cancer than in patients with HR+ and HER2− breast cancer (7, 43). Future studies should enroll more patients to assess the effects of standard treatment, so that the prediction model could be better trained. In addition, the model itself should also be further optimized with better engineering techniques, thereby improving the overall performance of RMM. The performance of RMM may also be investigated in patients of different ethnic populations in the future.

In conclusion, the present preliminary study suggested a possibility that RMM provided a potential tool to develop a model for predicting pCR to NAC in breast cancer. With further clinical research, a prediction model may be developed with the radiomics combining multiparametric MRI and clinical information.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: K. Wang, Zaiyi Liu, J. Tian
Development of methodology: Zaiyi Liu
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Zhenyu Liu, Z. Li, J. Qu, R. Zhang, H. Jiang, H. Li, Q. Xiong, Y. Ding, X. Zhao, K. Wang, Zaiyi Liu
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Zhenyu Liu, X. Zhou, L. Li, K. Sun, Z. Tang, Zaiyi Liu, J. Tian

Writing, review, and/or revision of the manuscript: Zhenyu Liu, X. Zhou, L. Li, K. Wang, Zaiyi Liu, J. Tian
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Tian
Study supervision: Zaiyi Liu

Acknowledgments

This paper is supported by the National Natural Science Foundation of China (Grant Nos. 81772012, 81771912, 81871513, and 81227901), the Beijing Natural Science Foundation under Grant No. 7182109, the National Key Research and Development Plan of China under Grant Nos. 2017YFA0205200 and 2017YFC1309100, and the Chinese Academy of Sciences under Grant No. GJJSTD20170004. The authors would like to acknowledge the instrumental and technical support of multimodal biomedical imaging experimental platform, Institute of Automation, Chinese Academy of Sciences.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 2, 2018; revised December 16, 2018; accepted March 5, 2019; published first March 6, 2019.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7–30.
2. Gradishar WJ, Anderson BO, Balassanian R, Blair SL, Burstein HJ, Cyr A, et al. Breast cancer, version 4.2017. NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2018;16:310–20.
3. Thompson AM, Moulder-Thompson SL. Neoadjuvant treatment of breast cancer. Ann Oncol 2012;23:231–6.
4. Derks MGM, van de Velde CJH. Neoadjuvant chemotherapy in breast cancer: more than just downsizing. Lancet Oncol 2018;19:2–3.
5. Cortazar P, Zhang LJ, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014;384:164–72.
6. Alberro JA, Ballester B, Deulofeu P, Fabregas R, Fraile M, Gubern JM, et al. Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. Lancet Oncol 2018;19:27–39.
7. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. J Clin Oncol 2012;30:1796–804.
8. Tran WT, Gangeh MJ, Sannachi L, Chin L, Watkins E, Bruni SG, et al. Predicting breast cancer response to neoadjuvant chemotherapy using pretreatment diffuse optical spectroscopic texture analysis. Brit J Cancer 2017;116:1329–39.
9. Dialani V, Chadashvili T, Slanetz PJ. Role of imaging in neoadjuvant therapy for breast cancer. Ann Surg Oncol 2015;22:1416–24.
10. Li H, Yao L, Jin P, Hu L, Li X, Guo T, et al. MRI and PET/CT for evaluation of the pathological response to neoadjuvant chemotherapy in breast cancer: a systematic review and meta-analysis. Breast 2018;40:106–15.
11. Weber JJ, Jochelson MS, Eaton A, Zabor EC, Barrio AV, Gemignani ML, et al. MRI and prediction of pathologic complete response in the breast and axilla after neoadjuvant chemotherapy for breast cancer. J Am Coll Surg 2017;225:740–6.
12. Marinovich ML, Houssami N, Macaskill P, Sardanelli F, Irwig L, Mamounas EP, et al. Meta-analysis of magnetic resonance imaging in detecting residual breast cancer after neoadjuvant therapy. Jnci-J Natl Cancer I 2013;105:321–33.
13. Park SH, Moon WK, Cho N, Song IC, Chang JM, Park IA, et al. Diffusion-weighted MR imaging: pretreatment prediction of response to neoadjuvant chemotherapy in patients with breast cancer. Radiology 2010;257:56–63.
14. Park SH, Moon WK, Cho N, Chang JM, Im SA, Park IA, et al. Comparison of diffusion-weighted MR imaging and FDG PET/CT to predict pathological complete response to neoadjuvant chemotherapy in patients with breast cancer. Eur Radiol 2012;22:18–25.
15. Partridge SC, Zhang Z, Newitt DC, Gibbs JE, Chenevert TL, Rosen MA, et al. Diffusion-weighted MRI findings predict pathologic response in neoadjuvant treatment of breast cancer: the ACRIN 6698 Multicenter Trial. Radiology 2018:180273.
16. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.
17. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563–77.
18. Aerts HJ. The potential of radiomic-based phenotyping in precisionmedicine a review. JAMA Oncol 2016;2:1636–42.
19. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. Theranostics 2019;9:1303–22.
20. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016;34:2157–64.
21. Guo J, Liu Z, Shen C, Li Z, Yan F, Tian J, et al. MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation. Eur Radiol 2018;28:3872–81.
22. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
23. Liu Z, Wang Y, Liu X, Du Y, Tang Z, Wang K, et al. Radiomics analysis allows for precise prediction of epilepsy in patients with low-grade gliomas. Neuroimage Clin 2018;19:271–8.
24. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
25. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuze S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol 2017;28:1191–206.
26. Liu ZY, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang ZC, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Clin Cancer Res 2017;23:7253–62.
27. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. Lancet Oncol 2018;19:1180–91.
28. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. Breast Cancer Res 2017;19:57.
29. Chamming's F, Ueno Y, Ferre R, Kao E, Jannot AS, Chong J, et al. Features from computerized texture analysis of breast cancers at pretreatment MR imaging are associated with response to neoadjuvant chemotherapy. Radiology 2018;286:412–20.
30. Fujii T, Kogawa T, Dong W, Sahin AA, Moulder S, Litton JK, et al. Revisiting the definition of estrogen receptor positivity in HER2-negative primary breast cancer. Ann Oncol 2017;28:2420–8.
31. Mazouni C, Peintinger F, Wan-Kau S, Andre F, Gonzalez-Angulo AM, Symmans WF, et al. Residual ductal carcinoma in situ in patients with complete eradication of invasive breast cancer after neoadjuvant chemotherapy does not adversely affect patient outcome. J Clin Oncol 2007;25:2650–5.
32. Symmans WF, Peintinger F, Hatzis C, Rajan R, Kuerer H, Valero V, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol 2007;25:4414–22.
33. Sikov WM, Berry DA, Perou CM, Singh B, Cirrincione CT, Tolaney SM, et al. Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and

cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (Alliance). J Clin Oncol 2015; 33:13–21.
34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
35. Kursa MB, Rudnicki WR. Feature selection with the boruta package. J Stat Softw 2010;36:1–13.
36. Sauerbrei W, Boulesteix AL, Binder H. Stability investigations of multivariable regression models derived from low- and high-dimensional data. J Biopharm Stat 2011;21:1206–31.
37. Collins GS, Reitsma JB, Altman DG, Moons KGM, members of the TRIPOD group. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur Urol 2015;67:1142–51.
38. Delong ER, Delong DM, Clarkepearson DI. Comparing the areas under 2 or more correlated receiver operating characteristic curves - a nonparametric approach. Biometrics 1988;44:837–45.
39. Fan M, Wu GL, Cheng H, Zhang J, Shao GL, Li LH. Radiomic analysis of DCE-MRI for prediction of response to neoadjuvant chemotherapy in breast cancer patients. Eur J Radiol 2017;94:140–7.
40. deSouza NM. Diffusion-weighted MRI in multicenter trials of breast cancer: a useful measure of tumor response? Radiology 2018;289:628–9.
41. Nie K, Shi LM, Chen Q, Hu X, Jabbour SK, Yue N, et al. Rectal cancer: assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. Clin Cancer Res 2016;22:5256–64.
42. Price ER, Wong J, Mukhtar R, Hylton N, Esserman LJ. How to use magnetic resonance imaging following neoadjuvant chemotherapy in locally advanced breast cancer. World J Clin Cases 2015;3:607–13.
43. Gianni L, Eiermann W, Semiglazov V, Manikhas A, Lluch A, Tjulandin S, et al. Neoadjuvant chemotherapy with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial): a randomised controlled superiority trial with a parallel HER2-negative cohort. Lancet 2010;375:377–84.