

Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer

Zhenyu Liu^{1,2}, Xiao-Yan Zhang², Yan-Jie Shi², Lin Wang³, Hai-Tao Zhu², Zhenchao Tang¹, Shuo Wang¹, Xiao-Ting Li², Jie Tian^{1,4}, and Ying-Shi Sun²



Abstract

Purpose: To develop and validate a radiomics model for evaluating pathologic complete response (pCR) to neoadjuvant chemoradiotherapy in patients with locally advanced rectal cancer (LARC).

Experimental Design: We enrolled 222 patients (152 in the primary cohort and 70 in the validation cohort) with clinicopathologically confirmed LARC who received chemoradiotherapy before surgery. All patients underwent T2-weighted and diffusion-weighted imaging before and after chemoradiotherapy; 2,252 radiomic features were extracted from each patient before and after treatment imaging. The two-sample *t* test and the least absolute shrinkage and selection operator regression were used for feature selection, whereupon a radiomics signature was built with support vector machines. Multivariable logistic regression analysis was then used to develop a radiomics model incorporating the radiomics signature and independent clinicopathologic

risk factors. The performance of the radiomics model was assessed by its calibration, discrimination, and clinical usefulness with independent validation.

Results: The radiomics signature comprised 30 selected features and showed good discrimination performance in both the primary and validation cohorts. The individualized radiomics model, which incorporated the radiomics signature and tumor length, also showed good discrimination, with an area under the receiver operating characteristic curve of 0.9756 (95% confidence interval, 0.9185–0.9711) in the validation cohort, and good calibration. Decision curve analysis confirmed the clinical utility of the radiomics model.

Conclusions: Using pre- and posttreatment MRI data, we developed a radiomics model with excellent performance for individualized, noninvasive prediction of pCR. This model may be used to identify LARC patients who can omit surgery after chemoradiotherapy. *Clin Cancer Res*; 23(23); 7253–62. ©2017 AACR.

Introduction

More than 100,000 people worldwide are diagnosed with rectal cancer annually; 70% are locally advanced rectal cancers (LARC). The current standard treatment for LARC is neoadjuvant chemoradiotherapy followed by total mesorectal excision (TME; refs. 1–3). Approximately 15% to 27% of patients will show pathologic complete response (pCR) to chemoradiother-

apy (4, 5), which led some investigators to question the use of TME in patients who achieve pCR. Several previous studies suggested that such patients usually have excellent long-term outcomes without surgery (6–9), and that the "wait and see" management approach that avoids surgery and preserves organs is a valid option (10). However, pCR can only be confirmed by histopathologic examination of surgically resected specimens, and creating a noninvasive, validated method to safely and accurately identify pCR patients after chemoradiotherapy remains a major challenge.

Medical imaging can noninvasively evaluate the therapeutic responses to chemoradiotherapy. Several investigators have proposed methods to identify good responders to chemoradiotherapy using various pre- or posttreatment imaging data including fludeoxyglucose positron emission tomography (11), T2-weighted MRI (T2WI; ref. 12), dynamic contrast-enhanced MRI (13), and diffusion-weighted imaging (DWI; refs. 14, 15). Although these imaging methods have the potential to predict or evaluate responses to chemoradiotherapy, their accuracy in evaluating pCR is limited. In a recent small-sample study, radiomics analysis based on pretreatment multiparametric MRI performed well in predicting pCR after chemoradiotherapy, albeit without independent validation (16). These results suggested that pretreatment multiparametric MRI may be associated with response to chemoradiotherapy and that radiomics

¹CAS Key Laboratory of Molecular Imaging, Institute of Automation, Beijing, China. ²Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiology, Peking University Cancer Hospital and Institute, Beijing, China. ³Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Gastrointestinal Surgery, Peking University Cancer Hospital and Institute, Beijing, China. ⁴University of Chinese Academy of Sciences, Beijing, China.

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

Z. Liu, X.-Y. Zhang, and Y.-J. Shi contributed equally to this article.

Corresponding Authors: Ying-Shi Sun, Peking University Cancer Hospital and Institute, No. 52 Fu Cheng Road, Hai Dian District, Beijing 100142, China. Phone: 86-10-88196822; Fax: 86-10-88196822; E-mail: sys27@163.com; and Jie Tian, Phone: 86-10-82618465; Fax: 86-10-62527995; E-mail: jie.tian@ia.ac.cn

doi: 10.1158/1078-0432.CCR-17-1038

©2017 American Association for Cancer Research.

Translational Relevance

In the present study, we developed and validated a radiomics model for noninvasive, individualized evaluation of pathologic complete response (pCR) in patients with locally advanced rectal cancer (LARC) based on pre- and posttreatment MRI data. The model's validation data showed it to be robust in its ability to detect pCR. In addition to improved pCR detection, the model (and the derived nomogram that incorporates the radiomics signature and independent clinicopathologic risk factors) provides patients and doctors with an effective tool for evaluating chemoradiotherapeutic outcomes in patients with LARC and for determining further treatment plans.

analysis may greatly contribute to determining whether pCR has been achieved after chemoradiotherapy.

Radiomics, which is based on advanced pattern recognition tools, involves the extraction of a large number of quantitative features from digital images to determine relationships between such features and the underlying pathophysiology (17, 18). Radiomics analysis of large imaging datasets has been successfully employed in the field of oncology for noninvasively profiling tumor heterogeneity (19), and there is a growing interest within the field in devising maps that display the associations between tumor heterogeneity and imaging features (20). Recent advances in radiomics have enabled oncologists to deliver more personalized medical care that takes into account phenotypic subtypes (21), as well as to assess therapeutic responses using posttreatment imaging features (22, 23). A recent radiomics study in patients with colorectal cancer proposed a nomogram to predict lymph node metastasis (24); this further confirmed the clinical value of radiomics. To that end, a radiomics model for pCR detection could vastly improve treatment strategy planning. As pretreatment MRI is associated with responses to chemoradiotherapy while posttreatment MRI directly reflects the posttreatment status, a radiomics model combining pre- and posttreatment MRI data may potentially predict pCR with accuracy.

In the present study, we aimed to develop and validate a radiomics model for individualized pCR evaluation after chemoradiotherapy in patients with LARC. Consistent with clinical practice, our work combined pre- and posttreatment MRI data to noninvasively evaluate the outcomes of such patients and to select LARC patients for whom surgery can be avoided.

Materials and Methods

Patients

This retrospective study was approved by the ethics committee of Beijing Cancer Hospital; the informed consent requirement was waived. All procedures performed in studies involving human participants were in accordance with the 1964 Helsinki declaration and its later amendments. A total of 222 patients who underwent surgical treatment between July 2010 and June 2015 were consecutively included in this study according to the following inclusion and exclusion criteria (Supplementary Fig. S1). The inclusion criteria were (i) biopsy-proven primary rectal adenocarcinoma; (ii) locally advanced

disease determined by pretreatment MRI ($\geq T3$, and/or positive nodal status); (iii) received complete neoadjuvant chemoradiotherapy and no treatment has been done before; (iv) TME surgery was performed after completion of neoadjuvant chemoradiotherapy, after which pCR was confirmed by postoperative pathologic examination; and (v) pre- and posttreatment MRI data obtained using the same 3-T MR scanner, including DWI and high-resolution T2WI. The exclusion criteria were (i) not completing neoadjuvant chemoradiotherapy; (ii) not undergoing surgery at our hospital, or pCR was not assessed; (iii) lack of DWI or high-resolution T2WI data; (iv) insufficient MRI quality to obtain measurements (e.g., owing to motion artifacts); (v) mucinous adenocarcinoma detected on pathologic examination after TME; and (vi) lack of presurgical carcinoembryonic antigen (CEA) and CA19-9 data. Patients were allocated to primary and validation cohorts according to the time of surgery in a 2:1 ratio; the first 152 patients were allocated to the primary cohort, and the subsequent 70 were allocated to the validation cohort. The clinical characteristics of all patients are shown in Table 1; the data analysis flowchart of the study is shown in Fig. 1.

Neoadjuvant chemoradiotherapy treatment

All patients received preoperative chemoradiotherapy followed by TME. Intensity-modulated radiation therapy (IMRT) was administered using a Varian Rapidarc system (Varian Medical Systems). The IMRT regimen comprised 22 fractions of 2.3 Gy (gross tumor volume, GtV) and 1.9 Gy (clinical target volume, CtV). A total dose of 50.6 Gy (GtV)/41.8 Gy (CtV) was administered 5 times per week over a period of 30 days (25, 26). The GtV was defined as the volume of the primary tumor including the mesorectum. The CtV was defined as the primary tumor, mesorectal region, presacral region, mesorectal lymph nodes, lateral lymph nodes, internal iliac lymph node chain, and pelvic wall area. Capecitabine treatment was administered concurrently with IMRT at a dose of 825 mg/m² orally twice per day. TME-based surgery was recommended 8 weeks after completing chemoradiotherapy. Adjuvant chemotherapy was routinely recommended; capecitabine alone, mFOLFOX6, or CapeOx was prescribed at the discretion of the physician.

Pathologic assessment of response

Surgically resected specimens were histopathologically examined and analyzed by an experienced pathologist and were further reviewed by a dedicated gastrointestinal pathologist, both of whom were blinded to the MRI data; pCR was defined as the absence of viable tumor cells in the primary tumor and lymph nodes.

MRI data acquisition and retrieval procedure

All patients underwent MRIs at 2 time points: Within 1 week before the initiation of neoadjuvant chemoradiotherapy and within 1 week before surgery (defined as pre- and posttreatment MRI, respectively). All MRIs were performed with a 3.0-T MR scanner (Discovery 750; GE Healthcare) using an 8-channel phased array body coil in the supine position. To reduce colonic motility, 20 mg of scopolamine butylbromide was injected intramuscularly 30 minutes prior to MRI scanning. Patients were not required to undergo bowel preparation before the examination. All patients underwent a conventional rectal MRI protocol including DWI and axial, coronal, and

sagittal T2-weighted images. DWI images were obtained using single-shot echo-planar imaging with 2 b-factors (0 and 1,000 s/mm²), and repetition time (TR) = 2,800 ms, echo time (TE) = 70 ms, field of view (FOV) = 340 × 340 mm, matrix = 256 × 256, thickness = 4.0 mm, and gap = 1.0 mm. Apparent diffusion coefficient (ADC) maps were generated automatically and included both b values in a monoexponential decay model. High-resolution T2WI images were obtained using fast recovery fast spin echo with TR = 5,694 ms, TE = 110 ms, FOV = 180 × 180 mm, echo train length = 24, matrix = 288 × 256, thickness = 3.0 mm, and gap = 0.3 mm.

All MRI scans were retrieved from the picture archiving and communication system for further image feature extraction.

Tumor masking

Pre- and posttreatment MRIs were analyzed by two radiologists (Dr. Zhang, a radiologist with 10 years of experience in rectal cancer imaging, and Dr. Shi, who has 7 years of experience in rectal cancer imaging); both were blinded to the histopathology results. The regions of interest (ROI) were created manually via the itk-SNAP software (www.itksnap.org) using the T2WI and DWI data, including the whole tumor and excluding the intestinal lumen. ROIs of rectal tumors before and after therapy were manually drawn on each slice. Before chemoradiotherapy, ROIs were drawn along the contour of the tumor as visualized by T2WI (slightly high signal), containing the surrounding chords and burrs. ROIs were placed on the high signal intensity region on DWI (b-value of 1,000 s/mm²) on each slice. If a highly suspicious tumor signal was still noted on T2WI (slightly high signal) after chemoradiotherapy, the ROI delineating criteria were the same as those before chemoradiotherapy. If a low, mixed-intensity, or any other non-normal rectal wall signal was detected in the tumor bed on T2WI (abnormal signal), the ROIs were drawn with contouring of the abnormal signal region. In cases where no abnormal signals were detected on T2WI (iso-intensity signal compared with the normal rectal wall), the ROIs were placed on the primary tumor bed region determined by T2WI before chemoradiotherapy. Due to the higher resolution of DWI compared with ADC maps, ROIs were detected with a b-value of 1,000 s/mm² first and then copied to the corresponding ADC maps. If a highly suspicious tumor signal (high signal) was noted on DWI, the ROIs were placed over the high-signal region. In cases where no high signal was detected on DWI compared with the normal rectal wall, the ROIs were placed on the primary tumor bed region as determined by DWI before chemoradiotherapy. Care was taken to avoid the magnetic susceptibility artifact during DWI. If no tumor signals were noted on postchemoradiotherapy T2WI, then the ROI of DWI was outlined on the corresponding tumor bed region based on pretreatment images.

Radiomic feature extraction and statistical analysis

MRI scans for each patient were normalized with z-scores in order to obtain a standard normal distribution of image intensities. Next, three groups of imaging features were extracted from the normalized pre- and posttreatment T2WI and DWI data with manually segmented ROIs: (i) 4 statistical features, (ii) 43 voxel-intensity computational features, and (iii) 516 wavelet features. Group 1 consisted of quantified tumor intensity characteristics with first-order statistics calculated from the histogram of all tumor intensities. Group

2 comprised textual features based on the quantification of intratumoral heterogeneity (i.e., differences in texture observed within the tumor volume); these features were all calculated using two-dimensional analysis and averaged for all slices within the three-dimensional tumor volume. Group 3 incorporated the calculated textural features from the wavelet decompositions of the original images, thereby focusing on the various frequency scales and different feature orientations within the tumor volume. All of these features have generally been used in previous radiomics studies (19, 21, 24). The final set comprised of 563 features for each modality (T2WI and DWI) per MRI scan, resulting in a total of 2,252 radiomic features per patient. All feature-extracting algorithms were implemented using the Matlab software (Math Works Inc.); details are provided in the Supporting Information. In addition, all statistical analyses were conducted with MatlabR2014b (Math Works Inc.). The reported statistical significance levels are all two-sided, with the statistical significance set at 0.05.

Inter- and intraobserver reproducibility evaluation

Interobserver and intraobserver reproducibility of ROI detection and radiomic feature extraction was initially determined using the T2WI data of 80 consecutive patients undergoing investigation between July 2010 and May 2011 for ROI-based radiomic feature generation in a blinded fashion by Dr. Zhang and Dr. Shi. To assess intraobserver reproducibility, each reader repeated the generation of radiomic features twice within a 1-week period following the same procedure. Intraclass correlation coefficients were used to evaluate the intra- and interobserver agreement in terms of feature extraction; we interpreted a coefficient of 0.81 to 1.00 as almost perfect agreement, 0.61 to 0.80 as substantial agreement, 0.41 to 0.60 as moderate agreement, 0.21 to 0.40 as fair agreement, and 0 to 0.20 as poor or no agreement (27). Many radiomic features described the shape and size of the ROIs; therefore, these values could also be used to evaluate the overall inter- and intraobserver reproducibility of the ROIs.

Feature selection method

To reduce overfitting or any type of bias in our radiomics model, two feature selection steps were used. First, the best features based on univariate statistical tests (two-sample *t* test) between pCR and non-pCR groups in the primary cohort were selected. Second, regularized multivariate logistic regression with the least absolute shrinkage and selection operator (LASSO) penalty was applied to the data of the primary cohort. With a linear combination of the selected features weighted by their respective coefficients, a model was used to estimate the chemoradiotherapy outcomes based on the radiomic features. The model was defined as follows:

$$y = \sum_{j=1}^d \beta_j x_j + \beta_0 + \varepsilon$$

Where y is 1 for patients with pCR and 0 for non-pCR patients; d is the number of features used in the model; x_j ($j = 1, 2, \dots, d$) is the feature; β_j ($j = 0, 1, 2, \dots, d$) is the model parameter, and ε is the error term.

Using regularized regression to estimate the parameters of the model, feature selection (by forcing many parameters to

zero value) can be performed simultaneously. The aim of this approach is to minimize the cost function:

$$\sum_{i=1}^N \left[y_i - S \left(\sum_{j=1}^d \beta_j x_{ij} + \beta_0 \right) \right]^2 + \lambda \sum_{j=1}^d |\beta_j|$$

Where y_i is the outcome of patient i , N is the number of patients, S is the sigmoid function, x_{ij} is the j th feature of the i th patient, and λ is the regularization parameter. The sigmoid function S is defined as follows:

$$S(x) = \frac{1}{1 + e^{-x}}$$

with the LASSO penalty $\sum_{j=1}^d |\beta_j|$ applied, leading to sparse models by setting some parameters (β_j s) to zero. Features with greater contributions to the model are selected.

Radiomics signature construction

The support vector machine (SVM) method was used to discriminate whether a patient achieved pCR in this study. A radiomics score was calculated for each patient using an SVM model with linear kernel training based on the selected features. Leave-one-out cross-validation (LOOCV) was employed to determine the optimal value of the regularization parameter C using the primary cohort. The C value that maximized the area under the receiver operating characteristic curve (AUC) in the primary cohort was selected as the optimal regularization parameter. Specifically, we tested values of $C \in [0.01, 1]$ with a step size of 0.02. After C was selected, the radiomics score for each patient in the validation cohort was calculated using the SVM model. The AUC, classification accuracy, positive predictive value (PPV), and negative predictive value (NPV) were calculated as metrics to assess the quantitative discrimination performance of the radiomics signature in both the primary and validation cohorts.

Development of the individualized radiomics model

Multivariable logistic regression analysis was conducted with the following clinical information: age, sex, posttreatment CEA, posttreatment CA19-9, histologic grade, pre- and posttreatment tumor length, pre- and posttreatment tumor thickness obtained from oblique axial T2WI (TTOA), pre- and posttreatment invasion distance beyond the muscularis propria (IDBMP), pre- and posttreatment shortest distance between the mesorectal fascia and the outer edge of the tumor extension (SDBMT), pre- and posttreatment total number of the lymph nodes detected by DWI (NLN), pre- and posttreatment minor axis length of the largest lymph node (MALLN), and radiomics signature. Backward step-wise selection was applied using the likelihood ratio test with Akaike information criterion employed as the stopping rule.

Based on the multivariable logistic analysis of the aforementioned clinical parameters in the primary cohort, a radiomics model for pCR detection was constructed with the selected variates to provide a quantitative tool for clinical use.

Apparent performance and validation of the radiomics model

Calibration curves accompanied by the Hosmer–Lemeshow test were plotted to assess the radiomics model; a significance test statistic implied that the model was not perfectly calibrated

(28). Harrell's C-index, classification accuracy, PPV, and NPV were measured to quantify the model's discriminatory performance. The model was subjected to bootstrapping validation (1,000 bootstrapping resamples including feature selection, model construction, and performance estimation) to achieve a relatively corrected performance.

The performance of the radiomics model was then tested in the validation cohort. The multivariable logistic regression formula devised based on the primary cohort was applied to the patients in the validation cohort, and the total points were calculated for each. Logistic regression was then performed using the total points of each patient in the validation cohort; the performance of the model was derived using regression analysis.

Calibration curves were calculated to determine the agreement between the estimated probability of pCR and the actual outcomes (i.e., the pCR rate) in both the primary and validation cohorts. In the graph, the y axis represents the actual rate of pCR, while the x axis represents the calculated probability of pCR. The diagonal blue line represents a perfect diagnosis by an ideal model, and the pink line represents the performance of the radiomics model; a fit that is closer to the diagonal blue line represents better performance. The calibration curve was drawn by plotting \hat{P} on the x axis and $P_c = [1 + \exp - (\gamma_0 + \gamma_1 L)]^{-1}$ on the y axis, where P_c is the actual probability, $L = \text{logit}(\hat{P})$, \hat{P} is the diagnosed probability, γ_0 is the corrected intercept, and γ_1 is the slope estimate.

Clinical use

Decision curve analysis was conducted to determine the clinical usefulness of the radiomics model by quantifying the net benefits at different threshold probabilities in the validation dataset (29).

Results

Clinical characteristics

The clinical characteristics of the patients are summarized in Table 1 and Supplementary Table S1. There were no significant differences between the 2 cohorts in terms of pCR prevalence (17.11% and 17.14% in the primary and validation cohorts, respectively, $P = 0.567$). There were no significant differences in other clinical characteristics between the primary and validation cohorts except for pretreatment N stage and posttreatment tumor length (Supplementary Table S1). In addition, a few clinical characteristics were significantly different between the pCR and non-pCR groups (Table 1), including posttreatment T stage, posttreatment N stage, pretreatment IDBMP, pretreatment MALLN, posttreatment TTOA, posttreatment IDBMP, and posttreatment SDBMT; all of these characteristics were included in the pCR evaluation model.

Feature selection and radiomics signature construction

Satisfactory inter- and intraobserver reproducibility of radiomic feature extraction was achieved (Supporting Information).

To devise the radiomics signature, we first performed univariate analysis (two-sample t tests) using the primary cohort as a prefilter. To avoid eliminating highly discriminative features on multivariate analysis rather than on univariate analysis, more features than those that showed significant differences between the two groups were included as compensation. All

Table 1. Characteristics of patients in the primary and validation cohorts

Characteristic	Primary cohort		P	Validation cohort		P
	pCR	Non-pCR		pCR	Non-pCR	
Age, mean ± SD, years	59.96 ± 12.32	55.98 ± 8.971	0.057	64.08 ± 7.267	57.91 ± 10.753	0.062
Gender (%)			0.371			0.764
Male	13 (50%)	75 (59.52%)		9 (75%)	41 (70.69%)	
Female	13 (50%)	51 (40.48%)		3 (25%)	17 (29.31%)	
Posttreatment CRT CEA (%)			0.086			0.459
Normal	26 (100%)	108 (85.71%)		12 (100%)	51 (87.93%)	
Abnormal	0 (0%)	18 (14.29%)		0 (0%)	7 (12.07%)	
Posttreatment CRT CA19-9 (%)			0.971			1
Normal	25 (96.15%)	118 (93.65%)		12 (100%)	55 (94.83%)	
Abnormal	1 (3.85%)	8 (6.35%)		0 (0%)	3 (5.17%)	
Histologic grade (%)			0.381			0.167
I	0 (0%)	0 (0%)		0 (0%)	2 (3.45%)	
II	22 (84.62%)	113 (89.68%)		8 (66.67%)	48 (82.76%)	
III	2 (7.69%)	10 (7.94%)		2 (16.67%)	4 (6.90%)	
IV	2 (7.69%)	3 (2.38%)		1 (8.33%)	4 (6.90%)	
V	0 (0%)	0 (0%)		1 (8.33%)	0 (0%)	
Pretreatment T stage (%)			0.287			1
T0	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
T1	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
T2	6 (23.08%)	15 (11.90%)		2 (16.67%)	8 (13.80%)	
T3	19 (73.08%)	96 (76.19%)		10 (83.33%)	45 (77.59%)	
T4a	1 (3.85%)	6 (4.76%)		0 (0%)	3 (5.17%)	
T4b	0 (0%)	9 (7.14%)		0 (0%)	2 (3.45%)	
Pretreatment N stage (%)			0.299			0.439
N0	2 (7.69%)	9 (7.14%)		1 (8.33%)	2 (3.45%)	
N1a	2 (7.69%)	5 (3.97%)		2 (16.67%)	5 (8.62%)	
N1b	5 (19.23%)	28 (22.22%)		1 (8.33%)	5 (8.62%)	
N2a	3 (11.54%)	35 (27.78%)		3 (0.25%)	10 (17.24%)	
N2b	14 (53.85%)	49 (38.89%)		5 (41.67%)	36 (62.07%)	
Posttreatment T stage (%)			<0.001^a			<0.001^a
T0	26 (100%)	6 (4.76%)		12 (100%)	1 (1.72%)	
T1	0 (0%)	7 (5.56%)		0 (0%)	4 (6.90%)	
T2	0 (0%)	54 (42.86%)		0 (0%)	23 (39.66%)	
T3	0 (0%)	57 (45.24%)		0 (0%)	30 (51.72%)	
T4a	0 (0%)	1 (0.79%)		0 (0%)	0 (0%)	
T4b	0 (0%)	1 (0.79%)		0 (0%)	0 (0%)	
Posttreatment N stage (%)			0.020^a			0.295
N0	26 (100%)	86 (68.25%)		12 (100%)	37 (63.80%)	
N1a	0 (0%)	23 (18.25%)		0 (0%)	7 (12.07%)	
N1b	0 (0%)	11 (8.73%)		0 (0%)	8 (13.79%)	
N2a	0 (0%)	3 (2.38%)		0 (0%)	5 (8.62%)	
N2b	0 (0%)	3 (2.38%)		0 (0%)	1 (1.72%)	
Pretreatment TL	41.38 ± 15.118	45.94 ± 13.784	0.133	41.83 ± 12.372	45.53 ± 13.507	0.384
Pretreatment TTOA	14.27 ± 4.006	15.77 ± 5.141	0.163	15.50 ± 4.275	17.28 ± 5.486	0.295
Pretreatment IDBMP	2.65 ± 2.497	5.27 ± 4.384	<0.01^a	4.67 ± 7.215	6.17 ± 6.344	0.467
Pretreatment SDBMT	5.85 ± 5.951	4.04 ± 3.844	0.051	3.75 ± 2.768	4.66 ± 5.857	0.604
Pretreatment NLN	12.38 ± 5.838	12.45 ± 5.790	0.957	10.00 ± 4.880	12.53 ± 4.946	0.110
Pretreatment MALLLN	5.88 ± 3.548	6.59 ± 3.371	0.339	4.83 ± 1.642	7.16 ± 3.583	0.032^a
Posttreatment TL	27.42 ± 9.892	29.89 ± 10.211	0.262	22.25 ± 6.440	27.33 ± 9.267	0.076
Posttreatment TTOA	8.50 ± 2.997	9.71 ± 3.237	0.08	6.75 ± 2.006	9.71 ± 3.195	<0.01^a
Posttreatment IDBMP	0.85 ± 1.541	2.37 ± 3.209	0.019^a	1.50 ± 1.977	2.74 ± 3.832	0.280
Posttreatment SDBMT	7.62 ± 5.224	5.35 ± 4.670	0.029^a	7.33 ± 4.141	6.88 ± 6.644	0.821
Posttreatment NLN	7.23 ± 3.840	7.69 ± 3.914	0.585	5.67 ± 2.871	7.19 ± 3.390	0.152
Posttreatment MALLLN	3.54 ± 1.702	4.40 ± 3.124	0.177	2.92 ± 1.443	4.26 ± 2.475	0.075
Radiomics score (mean ± SD)	0.7017 ± 0.2687	0.0995 ± 0.1128	<0.01^a	0.6957 ± 0.2756	0.0815 ± 0.1041	<0.01^a

NOTE: Chi-square or Fisher exact tests, as appropriate, were used to compare the differences in categorical variables (gender, posttreatment CEA, posttreatment CA19-9, histologic grade, pre- and posttreatment T stage, and pre- and posttreatment N stage), whereas a two-sample *t* test was used to compare the differences in age, Radiomics score, pre- and posttreatment TL, pre- and posttreatment TTOA, pre- and posttreatment IDBMP, pre- and posttreatment SDBMT, pre- and posttreatment NLN, and pre- and posttreatment MALLLN. Laboratory analysis of CEA and CA 19-9 was done via routine blood tests within 1 week before surgery. The threshold value for CEA level was ≤ 5ng/mL and >5 ng/mL, and the threshold value for CA 19-9 level was ≤ 39 U/mL and >39 U/mL, according to the normal range used in clinics.

Abbreviations: CA19-9, carbohydrate antigen 19-9; TL, tumor's length.

^a*P* < 0.05.

features were sorted in increasing order of *P* values, and the top 676 features (30%) were included in the next step of the analysis with *P* < 0.0987. Next, 676 features were reduced to

30 potential predictors by applying regularized regression to the primary cohort with the LASSO penalty using LOOCV via minimum criteria (Fig. 2).

Downloaded from http://aacrjournals.org/clincoerres/article-pdf/23/23/7253/2041778/7253.pdf by guest on 13 June 2024

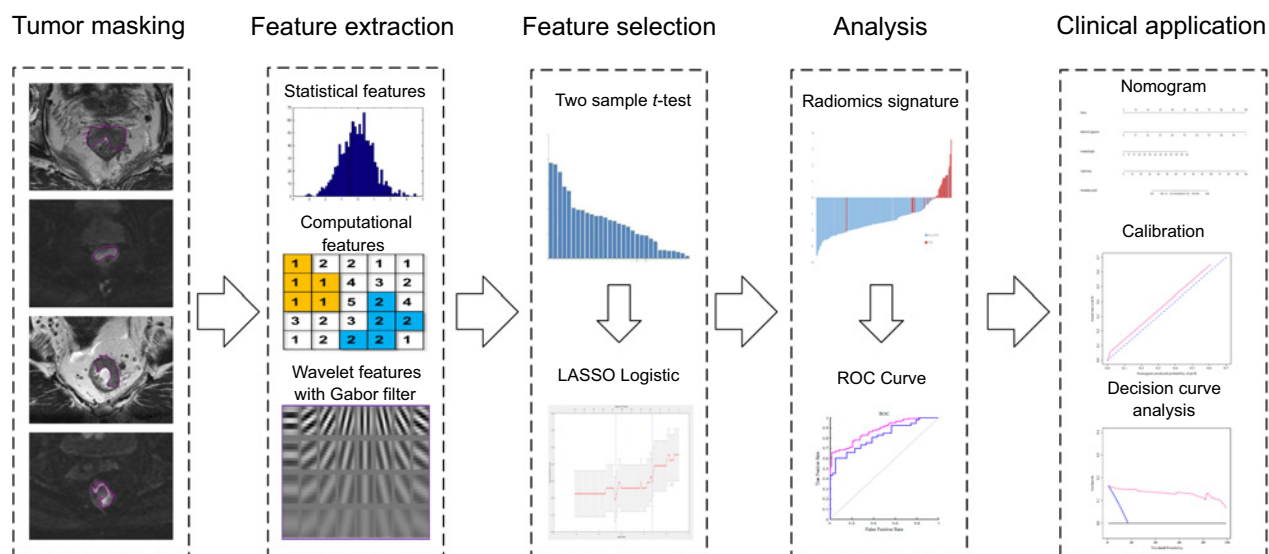


Figure 1. Flowchart of the study. With manually segmented tumor masks, we first extracted 2252 quantitative radiomic features from masked pre- and posttreatment T2-weighted imaging and diffusion-weighted imaging data; the general view of the feature extraction algorithm is shown. Next, two feature selection steps were applied on the extracted features with a two-sample *t* test and the least absolute shrinkage and selection operator (LASSO). Thereafter, a radiomics signature was constructed with the selected features using the linear kernel support vector machine model. Finally, the radiomics signature and clinical factors were incorporated into a nomogram for individual evaluation. ROC, receiver operating characteristic.

Next, an SVM model with a linear kernel was constructed using the selected features based on the primary cohort. The best regularization parameter ($C = 0.05$) was determined by LOOCV. The resultant coefficients of the features used in calculating the radiomics score are shown in Supplementary Table S2. The distributions of the radiomics scores and outcomes of chemor-

adiotherapy for each patient in the primary and validation cohorts are shown in Fig. 3.

Diagnostic validation of the radiomics signature

There was a significant difference in radiomics scores between pCR and non-pCR patients in the primary cohort ($P < 0.01$);

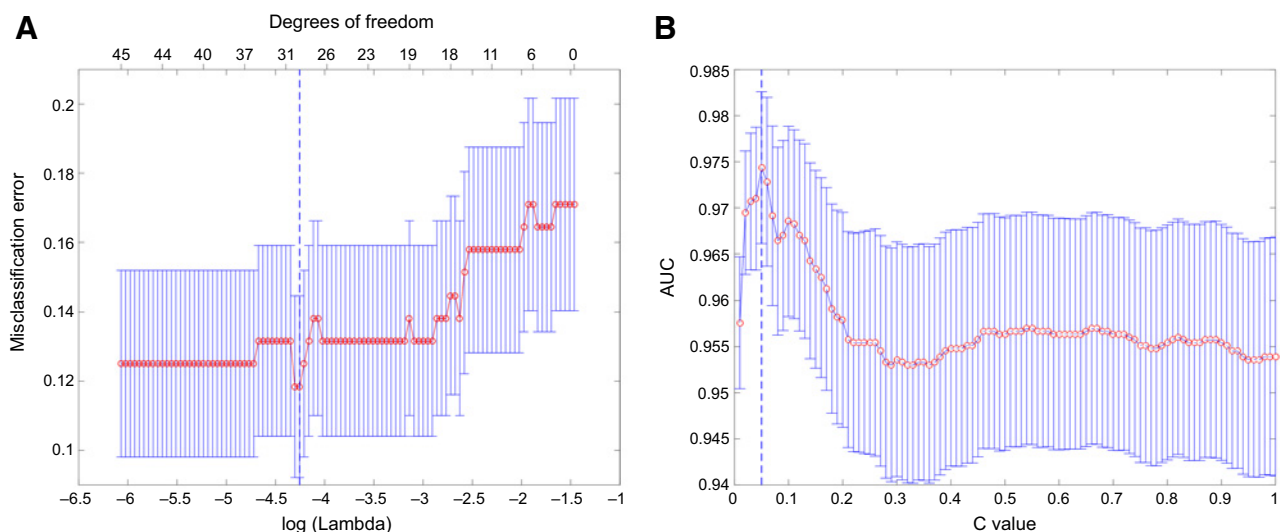


Figure 2. Tuning parameter selection for feature selection (λ) and support vector machine model construction (C). **A**, Feature selection with least absolute shrinkage and selection operator using LOOCV via minimum criteria. The misclassification error was plotted versus $\log(\lambda)$. A dotted vertical line was drawn at the optimal value by using the minimum criteria. A λ value of 0.014189 was chosen according to the LOOCV. **B**, The optimal value regularization parameter C selection for the SVM model construction using LOOCV. The AUC was plotted vs. C . A dotted vertical line was drawn at the optimal value. The C value of 0.05 was chosen according to the LOOCV.

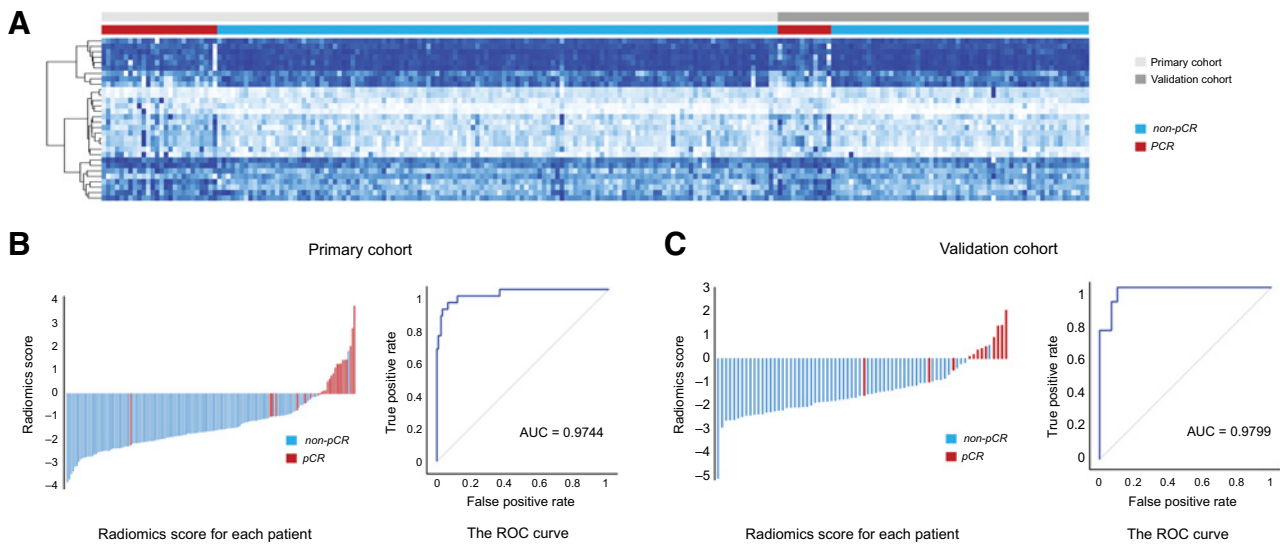


Figure 3. Radiomic features and performance of the radiomics signature. **A**, Heatmap of 30 selected radiomic features. Each row corresponds to 1 radiomic feature, and each column corresponds to 1 patient (separately grouped for the primary vs. validation cohort and the pCR vs. non-pCR group). **B**, The radiomics score for each patient and ROC curve in the primary cohort. **C**, The radiomics score for each patient and ROC curve in the validation cohort.

the same was true in the validation cohort ($P < 0.01$). The radiomics signature yielded an AUC of 0.9744 [95% confidence interval (CI), 0.9642–0.9756] and a classification accuracy of 94.08% (95% CI, 93.19–94.79%) in the primary cohort, and an AUC of 0.9799 (95% CI, 0.9780–0.9840) and a classification accuracy of 94.29% (95% CI, 94.21–95.61%) in the validation cohort. More importantly, the radiomics signature achieved a PPV of 86.96% (95% CI, 84.84–90.40%) in the primary cohort and 90.00% (95% CI, 89.60–99.40%) in the validation cohort. Detailed information on radiomics signature performance is shown in Table 2.

Development, performance, and validation of the individualized radiomics nomogram

The radiomics signature and posttreatment tumor length were identified as independent factors predicting pCR (Supplementary Tables S3 and S4). The model that incorporated these independent predictors was developed and presented as a nomogram (Fig. 4).

The calibration curve of the radiomics model estimating the probability of pCR demonstrated good agreement in the primary cohort (Fig. 4). The Hosmer–Lemeshow test yielded a nonsignificant statistic ($P = 0.9609$), suggesting no departure from the perfect fit. The C-index for the radiomics model was 0.9799 (95% CI, 0.9517–1) within the primary cohort.

Good performance was also observed for the probability of pCR in the validation cohort (Fig. 4). The Hosmer–Lemeshow test yielded nonsignificant statistics for the radiomics model ($P = 0.5416$). The radiomics model also achieved good discrimination performance with a C-index of 0.9756 (95% CI, 0.9417–1) and classification accuracy of 94.29% (95% CI, 91.85%–97.11%).

Clinical use

The decision curve analysis result for the radiomics model is shown in Fig. 4. The decision curve showed relatively good performance for the model in terms of clinical application. Although the probability of achieving pCR ranges from 0% to 100%, using the proposed radiomics model to detect pCR shows a greater advantage than either the scheme in which all patients are assumed to achieve pCR or the scheme in which no patients are.

Discussion

In the present study, we developed and validated a radiomics model that incorporated pre- and posttreatment MRI data for noninvasive, individualized prediction of pCR in patients with LARC. The easy-to-use nomogram facilitated noninvasive estimation of pCR. The proposed radiomics model performs well

Table 2. Performance of the radiomics signature and radiomics model

Metrics	Radiomics signature		Radiomics model	
	Primary cohort	Validation cohort	Primary cohort	Validation cohort
Accuracy (95%)	94.08% (93.19%–94.79%)	94.29% (94.21%–95.61%)	96.05% (94.67%–97.59%)	94.29% (91.85%–97.11%)
AUC (95%)	0.9744 (0.9642–0.9756)	0.9799 (0.9780–0.9840)	0.9799 (0.9517–1)	0.9756 (0.9417–1)
PPV (95%)	86.96% (84.84%–90.40%)	90.00% (89.60%–99.40%)	91.67% (86.56%–97.06%)	90.00% (79.51%–99.12%)
NPV (95%)	95.35% (94.39%–95.81%)	95.00% (94.87%–95.17%)	96.88% (95.42%–98.34%)	95.00% (92.19%–97.65%)

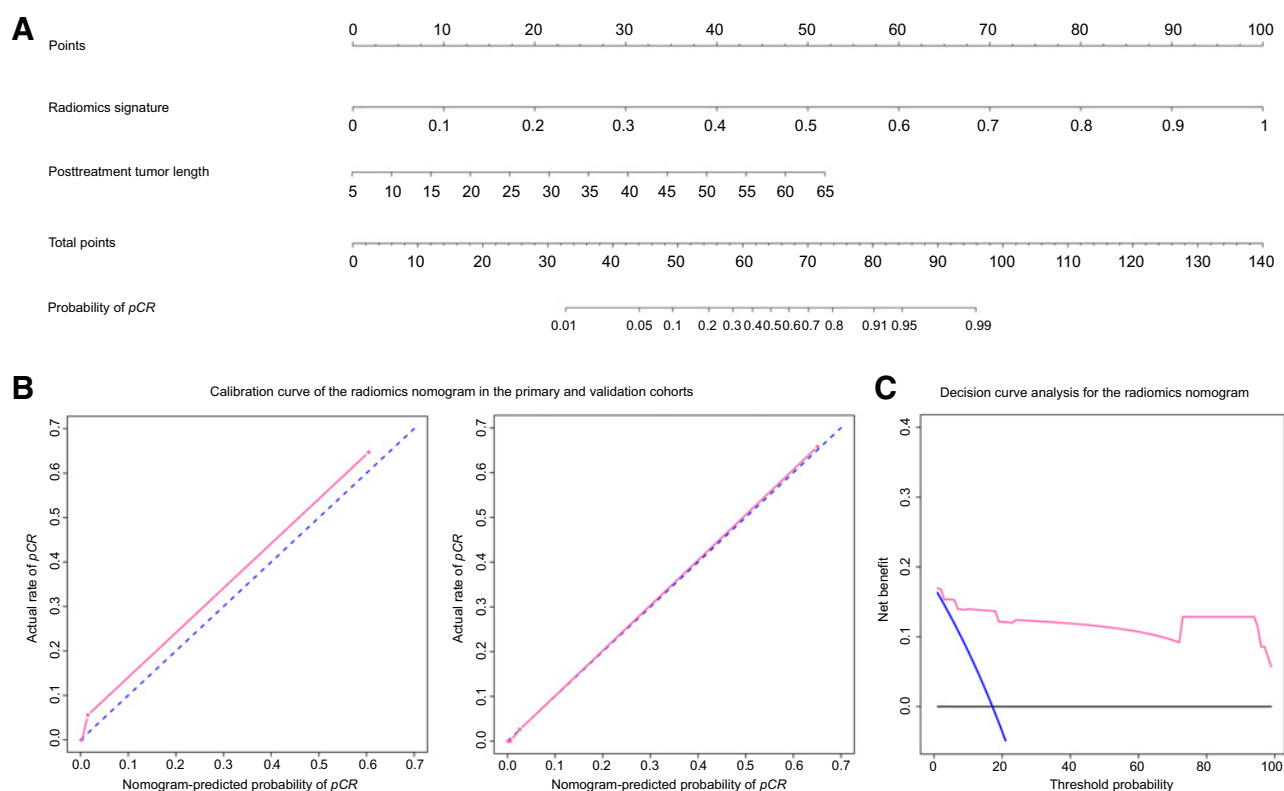


Figure 4.

Nomogram developed with the radiomics model and calibration curves, as well as the decision curve derived from the radiomics model. **A**, The developed radiomics nomogram. **B**, Calibration curves of the radiomics model in the primary and validation cohorts. Calibration curves depict the calibration of each model in terms of the agreement between the predicted probability of pCR and actual outcomes of the pCR rate. The y axis represents the actual rate of pCR. The x axis represents the predicted probability of pCR. The diagonal blue line represents a perfect prediction by an ideal model. The pink line represents the performance of the radiomics model, of which a closer fit to the diagonal blue line represents a better prediction. **C**, Decision curve analysis for the radiomics model. The y axis measures the net benefit. The pink line represents the radiomics model. The blue line represents the assumption that all patients showed pCR. The black line represents the assumption that no patients showed pCR.

and thereby provides an effective tool for clinical decision-making.

The accurate detection of pCR using visual judgment (conventional MRI) remains challenging in clinical settings. Methods using multimodality MRI (e.g., combining DWI and conventional MRI; refs. 14, 28–30) or positron emission tomography/CT (31) may also perform well; however, their levels of accuracy are not clinically reliable. Radiomics analysis integrates many high-dimensional imaging features used to evaluate pCR that are difficult to detect visually. Our proposed radiomics model based on these imaging features performed better than previously reported methods and can therefore be useful in clinical decision-making as it provides radiologists and oncologists with a potential quantitative tool for individualized pCR prediction.

To use our proposed radiomics model, radiologists must first delineate the ROI on pre- and posttreatment MRI scans (T2WI and DWI), after which the model allows for the calculation of the probability of pCR for each individual patient. Oncologists can then consider various factors, including the calculated probability of pCR and other retrievable clinical information, as well as their own clinical experience and the patient's opinion, to make a comprehensive

judgment on whether a wait-and-see treatment approach is warranted.

The radiomics model combined pre- and posttreatment T2WI and DWI data of patients with LARC and demonstrated adequate discrimination in both the primary and validation cohorts. There were significant differences in pretreatment N stage and posttreatment tumor length between these two cohorts. Nonetheless, the proposed radiomics model still performed appropriately and was well-calibrated. The results suggest that the radiomics model is robust in its evaluation of pCR and can be used in the clinical setting. Two recent studies investigated the pCR prediction capability of texture or radiomic features with DWI and multiparametric MRI without independent validation; they derived AUCs less than 0.9 (15, 16), which was a much lower value than the independent validation results obtained in our study. Specifically, our proposed radiomics model achieved a relatively high NPV and PPV in both the primary and validation cohorts. The high NPV indicated that the non-pCR evaluation of the proposed model was reliable. Thus, surgeons may potentially forgo colonoscopies or other examinations meant to confirm the absence of residual lesions in non-pCR patients and can thereby avoid excessive treatments that would ensue in the event that a pCR patient is incorrectly

judged to be a non-pCR patient. Conversely, the high PPV suggests that our model can satisfactorily enable surgeons to screen out pCR patients, allowing for a "watch and wait" approach. Patients who were designated pCR using our model had a relatively high probability of achieving true pCR.

An explanation for the robustness and improved performance of our radiomics model is the use of ADC maps derived from DWI. We obtained 2,252 features from the MRI data of each patient; after feature selection with a two-sample *t* test and LASSO logistic regression, 30 potential predictors were further analyzed. Only one pretreatment T2WI feature was selected with LASSO for the construction of the radiomics signature, suggesting that T2WI was not a good option for pCR assessment after chemoradiotherapy. Several studies have shown the difficulty of identifying pCR using the morphological features exposed by T2WI (14, 30, 32). As such morphological features reflect only limited information about residual cancer cells after chemoradiotherapy, DWI may provide more useful details. As a functional imaging technique, DWI showed strong potential in detecting subtle cancer cell remnants (28) and added valuable information regarding the responses to chemoradiotherapy in patients with LARC (15, 29). The use of DWI may improve the performance and confidence of radiologists in selecting patients with pCR after chemoradiotherapy compared with conventional T2WI (14, 15, 30).

Another reason for the robustness of our model was the combination of pre- and posttreatment MRI data during analysis; this differed from what was done in a recent study (16). This combination is reflective of clinical practice and encompasses the entire diagnosis and treatment process. Most importantly, posttreatment MRI data represent the current status of the tumor after chemoradiotherapy; the data or information contained in posttreatment MRI scans correspond more closely to the surgical pathology. Hence, including posttreatment MRI data improves the model's reliability in detecting pCR. Our results showed that 19 of the 30 selected radiomic features were from posttreatment MRI data.

Moreover, the use of high-dimensional features also contributed to the performance of the model. Previous studies generally used low-dimensional information to evaluate the responses to chemoradiotherapy (23, 28, 30). However, in the present study, 90% ($n = 27$) of the key features in the radiomics model were Gabor filtered wavelet features. Although the morphological and textural features of tumors can easily be discerned, high-dimensional features are challenging to decipher with the naked eye (Supplementary Fig. S2), and ensuring that every clinician achieves a high level of expertise in gleaning detailed information from imaging features remains a significant obstacle. However, high-dimensional features hold more detailed information about the cancer and are more sensitive when assessing pCR, as was also demonstrated in a recent study (16). Hence, by incorporating these high-dimensional imaging features, a radiomics-based model can assist doctors in accurately identifying patients with pCR for whom a "wait and see" approach may be the most appropriate.

We used a nomogram as an individualized tool for pCR detection and assessed whether the radiomics nomogram-based decisions could benefit patients. Decision curve analysis was applied to examine the clinical consequences

based on threshold probability, from which a net benefit (defined as the proportion of true positives minus the proportion of false positives, weighted by the relative harm of false positive and false negative results) could be derived (26, 33). The decision curve analysis proved that, given a threshold probability ranging from 0% to 100%, using the radiomics model to detect chemoradiotherapy outcomes provides a greater advantage than either the treat-all or treat-none scheme.

The use of the radiomics model not only provided an individualized tool for establishing a treatment plan, but also incorporated the radiomics signature and other clinical risk factors (age, sex, posttreatment CEA, posttreatment CA19-9, pre- and posttreatment tumor length, pre- and posttreatment TTOA, pre- and posttreatment IDBMP, pre- and posttreatment SDBMT, pre- and posttreatment NLN, and pre- and posttreatment MALLN). The constructed radiomics model comprised of the radiomics signature and posttreatment tumor length. To the best of our knowledge, the posttreatment tumor length has never been proposed for pCR detection. Although its β value in the multivariate regression model was minuscule compared with that of the radiomics signature, posttreatment tumor length may provide complementary information for precise evaluation of pCR using the radiomics model. The potential association between posttreatment tumor length and pCR could be further investigated in future studies.

There were some limitations of the study. First, the sample size of patients with pCR was small relative to the entire cohort. Second, all the patients were from a single center. Although we categorized the patients into independent primary and validation cohorts according to their surgery dates, the model may perform differently if multicenter datasets with different parameters are used. A much larger dataset from multiple centers, with a considerably large sample of patients with pCR, ought to be investigated to validate the robustness and reproducibility of our proposed radiomics model.

Conclusion

We propose a validated and easy-to-use radiomics model based on pre- and posttreatment MRI data for the individualized detection of pCR in patients with LARC. This model provides a noninvasive and convenient method to guide treatment planning in patients with LARC after they have undergone chemoradiotherapy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: J. Tian, Y.-S. Sun

Development of methodology: Z. Liu, X.-Y. Zhang, Y.-J. Shi, Z. Tang, S. Wang, Y.-S. Sun

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): X.-Y. Zhang, Y.-J. Shi, L. Wang

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Z. Liu, X.-Y. Zhang, Y.-J. Shi, H.-T. Zhu, Z. Tang, X.-T. Li, Y.-S. Sun

Writing, review, and/or revision of the manuscript: Z. Liu, X.-Y. Zhang, Y.-J. Shi, H.-T. Zhu, X.-T. Li, J. Tian, Y.-S. Sun

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Z. Liu, X.-Y. Zhang, Y.-J. Shi, X.-T. Li, Y.-S. Sun
Study supervision: Z. Liu, J. Tian, Y.-S. Sun

Grant Support

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81471640, 81501549, 81227901, 81501621, and 61231004), the Key Research Program of the Chinese Academy of Sciences (Grant No. KGZD-EW-T03), the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (Grant No. KFJ-SW-STS-160), and the Beijing Municipal Science & Technology Commission 'Capital Clinical Research Special Fund' (Nos. Z171100001017102 and Z151100004015105).

References

- Kapiteijn E, Marijnen CA, Nagtegaal ID, Putter H, Steup WH, Wiggers T, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *N Engl J Med* 2001;345:638–46.
- van de Velde CJ, Boelens PG, Borras JM, Coebergh JW, Cervantes A, Blomqvist L, et al. EURECCA colorectal: multidisciplinary management: European consensus conference colon & rectum. *Eur J Cancer* 2014;50:e1–1.e34.
- van Gijn W, Marijnen CA, Nagtegaal ID, Kranenbarg EM, Putter H, Wiggers T, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. *Lancet Oncol* 2011;12:575–82.
- Sanghera P, Wong DW, McConkey CC, Geh JI, Hartley A. Chemoradiotherapy for rectal cancer: an updated analysis of factors affecting pathological response. *Clin Oncol* 2008;20:176–83.
- Maas M, Nelemans PJ, Valentini V, Das P, Rodel C, Kuo LJ, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol* 2010;11:835–44.
- Habr-Gama A, Perez RO, Proscurshim I, Campos FG, Nadalin W, Kiss D, et al. Patterns of failure and survival for nonoperative treatment of stage c0 distal rectal cancer following neoadjuvant chemoradiation therapy. *J Gastrointest Surg* 2006;10:1319–28.
- Borschitz T, Wachtlin D, Mohler M, Schmidberger H, Junginger T. Neoadjuvant chemoradiation and local excision for T2-3 rectal cancer. *Ann Surg Oncol* 2008;15:712–20.
- Maas M, Beets-Tan RGH, Lambregts DMJ, Lammering G, Nelemans PJ, Engelen SME, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol* 2011;29:4633–40.
- Renahan AG, Malcomson L, Emsley R, Gollins S, Maw A, Myint AS, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *Lancet Oncol* 2016;17:174–83.
- Marijnen CAM. Organ preservation in rectal cancer: have all questions been answered? *Lancet Oncol* 2015;16:E13–22.
- Janssen MHM, Ollers MC, Riedl RG, van den Bogaard J, Buijssen J, van Stiphout RGPM, et al. Accurate prediction of pathological rectal tumor response after two weeks of preoperative radiochemotherapy using (18)F-fluorodeoxyglucose-positron emission tomography-computed tomography imaging. *Int J Radiat Oncol* 2010;77:392–9.
- Engelen SME, Beets-Tan RGH, Lahaye MJ, Lammering G, Jansen RLH, van Dam RM, et al. MRI after chemoradiotherapy of rectal cancer: a useful tool to select patients for local excision. *Dis Colon Rectum* 2010;53:979–86.
- Intven M, Reerink O, Philipppens MEP. Dynamic contrast enhanced MR imaging for rectal cancer response assessment after neo-adjuvant chemoradiation. *J Magn Reson Imaging* 2015;41:1646–53.
- Lambregts DMJ, Vandecaveye V, Barbaro B, Bakers FCH, Lambrecht M, Maas M, et al. Diffusion-weighted MRI for selection of complete responders after chemoradiation for locally advanced rectal cancer: a multicenter study. *Ann Surg Oncol* 2011;18:2224–31.
- Blazic IM, Lilic GB, Gajic MM. Quantitative assessment of rectal cancer response to neoadjuvant combined chemotherapy and radiation therapy: comparison of three methods of positioning region of interest for ADC measurements at diffusion-weighted MR imaging. 2017;282:418–28.
- Nie K, Shi L, Chen Q, Hu X, Jabbour SK, Yue NJ, et al. Rectal cancer: assessment of neoadjuvant chemo-radiation outcome based on radiomics of multi-parametric MRI. *Clin Cancer Res* 2016;22:5256–64.
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441–6.
- Kumar V, Gu YH, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30:1234–48.
- Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5:4006.
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;278:563–77.
- Itakura H, Achrol AS, Mitchell LA, Loya JJ, Liu T, Westbroek EM, et al. Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. *Sci Transl Med* 2015;7:303ra138.
- Altazi B, Fernandez D, Zhang C, Biagioli M, Moros E. Prediction of cervical cancer treatment response using radiomics features based on F18-FDG uptake in PET images. *Med Phys* 2015;42:3326.
- Kickingereder P, Gotz M, Muschelli J, Wick A, Neuberger U, Shinohara RT, et al. Large-scale radiomic profiling of recurrent glioblastoma identifies an imaging predictor for stratifying anti-angiogenic treatment response. *Clin Cancer Res* 2016;22:5765–71.
- 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.
- Wang L, Li ZY, Li ZW, Li YH, Sun YS, Ji JF, et al. Efficacy and safety of neoadjuvant intensity-modulated radiotherapy with concurrent capecitabine for locally advanced rectal cancer. *Dis Colon Rectum* 2015;58:186–92.
- Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* 2015;16:E173–80.
- Chang CC, Lin CJ. LIBSVM: a library for support vector machines. *Acm T Intel Syst Tec* 2011;2.
- van der Paardt MP, Zagers MB, Beets-Tan RGH, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology* 2013;269:101–12.
- Kim SH, Lee JM, Hong SH, Kim GH, Lee JY, Han JK, et al. Locally advanced rectal cancer: added value of diffusion-weighted MR imaging in the evaluation of tumor response to neoadjuvant chemo- and radiation therapy. *Radiology* 2009;253:116–25.
- Curvo-Semedo L, Lambregts DMJ, Maas M, Thywissen T, Mehnen RT, Lammering G, et al. Rectal cancer: assessment of complete response to preoperative combined radiation therapy with chemotherapy-conventional MR volumetry versus diffusion-weighted MR imaging. *Radiology* 2011;260:734–43.
- Song I, Kim SH, Lee SJ, Choi JY, Kim MJ, Rhim H. Value of diffusion-weighted imaging in the detection of viable tumour after neoadjuvant chemoradiation therapy in patients with locally advanced rectal cancer: comparison with T2 weighted and PET/CT imaging. *Br J Radiol* 2012;85:577–86.
- Franklin JM, Anderson EM, Gleeson FV. MRI features of the complete histopathological response of locally advanced rectal cancer to neoadjuvant chemoradiotherapy. *Clin Radiol* 2012;67:546–52.
- Collins GS, Reitsma JB, Altman DG, Moons KGM, Grp T. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *Eur Urol* 2015;67:1142–51.