Original Article

Magnetic resonance imaging-based radiomics in predicting the expression of Ki-67, p53, and epidermal growth factor receptor in rectal cancer

Qiying Li[1], Jinkai Liu[1], Weneng Li[1], Mingzhu Qiu[1], Xiaohua Zhuo[1], Qikui You[1], Shaohua Qiu[1], Qi Lin[1*], Yi Liu[2*]

[1]Longyan First Affiliated Hospital of Fujian Medical University, Longyan, China; [2]Liaoning Cancer Hospital and Institute, Shenyang, China

Contributions: (I) Conception and design: Q Li, Q Lin, Y Liu; (II) Administrative support: Q Lin; (III) Provision of study materials or patients: Q Li, W Li, M Qiu, X Zhuo, Q You; (IV) Collection and assembly of data: Q Li; (V) Data analysis and interpretation: J Liu, S Qiu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*These authors contributed equally to this work.

Correspondence to: Yi Liu, MD, PhD. Liaoning Cancer Hospital and Institute, No. 44 Xiaoheyan Load, Dadong District, Shenyang 110001, China. Email: liuyicmu@sina.cn; Qi Lin, MD, BD. Longyan First Affiliated Hospital of Fujian Medical University, No. 105 Jiuyibei Load, Xinluo District, Longyan 364000, China. Email: linqi_ly@126.com.

Background: The preoperative evaluation of the expression levels of Ki-67, p53, and epidermal growth factor receptor (EGFR) based on magnetic resonance imaging (MRI) of rectal cancer is necessary to facilitate individualized therapy. This study aimed to develop and validate radiomics models for the evaluation of the expression levels of Ki-67, p53, and EGFR of rectal cancer from preoperative MRI.

Methods: In this retrospective study, 124 patients (38 in the test group and 86 in the training group) with rectal cancer who underwent preoperative MRI and postoperative Ki-67, p53 and EGFR assay were included in Longyan First Affiliated Hospital of Fujian Medical University from June 2015 to October 2019. A total of 796 radiomics features were acquired from both diffusion-weighted imaging (DWI) and T2-weighted imaging (T2WI). Least absolute shrinkage and selection operator (LASSO) and the minimum redundancy maximum relevance (mRMR) were used to select the most predictive texture features, and then the radiomics score (Rad-score) models were derived to evaluate Ki-67, p53, and EGFR expression status based on the radiomics signature. The receiver operating characteristic (ROC) was used to assess the model's performance, and the reliability was verified via accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Results: The Rad-score evaluation of Ki-67 expression status yielded area under the curve (AUC) values of 0.91 [95% confidence interval (CI): 0.87–0.95] and 0.81 (95% CI: 0.66–0.96) in the training and test groups. The evaluation of p53 expression produced AUC values of 0.82 (95% CI: 0.77–0.88) and 0.80 (95% CI: 0.65–0.96). For evaluating EGFR expression status in both training and test groups, the AUC values were 0.86 (95% CI: 0.81–0.91) and 0.76 (95% CI: 0.58–0.93), respectively. While Rad-score of Ki-67 expression status in the training group obtained the top accuracy, sensitivity, specificity, and PPV with values of 0.85, 0.80, 0.92, and 0.93.

Conclusions: Preoperative MRI-based radiomics analysis has the ability to noninvasively assess the postoperative Ki-67, p53, and EGFR of rectal cancer.

Keywords: Radiomics signature; Ki-67; p53; epidermal growth factor receptor (EGFR)

Submitted Jun 07, 2024. Accepted for publication Sep 19, 2024. Published online Oct 29, 2024.
doi: 10.21037/jgo-24-220
View this article at: https://dx.doi.org/10.21037/jgo-24-220

Introduction

Colorectal cancer (CRC) ranks as the third highest in incidence among malignant diseases worldwide (1). Rectal cancer is a unique subgroup of CRC and a crucial contributor to cancer-related death.

Several immunohistochemical biomarkers have been identified (2) to evaluate biological behavior as a marker of tumor cell proliferation, such as the expression levels of Ki-67, p53, and epidermal growth factor receptor (EGFR). The expression of Ki-67 helps to enhance the proliferation of malignant tumors, and the Ki-67 index is regarded as a biomarker of tumor aggressiveness. The overexpression of Ki-67 in lesions was related to worse overall survival or disease-free survival (3). The p53 protein inhibits the advancement of solid and hematological tumors (4), which has been shown to be mutated in virtually all tumor entities. EGFR is a tumor oncogene that regulates a number of cell responses and is involved in cell differentiation, proliferation and apoptosis, and EGFR amplifications or overexpression could direct patients to targeted therapies (5). Hur *et al.* (6) found that low p53 expression and/or high p21 and Ki-67 expression were associated with a greater pathologic chemical response rate and that a scoring system based on biomarker expression levels had a favorable negative predictive value (NPV) and good sensitivity for predicting pathologic chemical response. However, the expression of Ki-67, p53, and EGFR of tumor can only be determined by using biopsy specimens before surgery. The invasive and incomplete sampling for detecting tumor immunohistochemical markers heightens the risk of several complications, including the risk of bleeding and tumor metastasis. And thus it is challenging for clinicians to make accurate judgments of the expression levels of Ki-67, p53, and EGFR. Therefore, it is crucial to find a non-invasive method for preoperatively assessing Ki-67, p53 and EGFR expression to guide the surgical or chemoradiotherapy strategy decision and monitor the tumor's progression or treatment response.

Medical imaging can comprehensively, noninvasively, and repeatedly evaluate tumor's characteristics in real time. In particular, magnetic resonance imaging (MRI) can be used to stage rectal cancer staging according to the imaging features of lesions before treatment (7,8). T2-weighted imaging (T2WI), a traditional MRI sequence, provides good tissue resolution and can intuitively assess the tumor's location and extent. Diffusion-weighted imaging (DWI) is a functional MRI technique that can be used to noninvasively analyze the diffusion process of water molecules in tissues *in vivo*. However, in clinical work, radiologists assess tumor characteristics, which include signal intensity, diameter, and boundaries, based on the naked eye and their individual experience in T2WI and DWI, which involves a degree of subjectivity.

Radiomics can reveal the potential relationship between medical images and microscopic pathophysiological features. It does this by extracting a large number of quantitative features from medical images for machine learning–based analysis, which can overcome the limitations of subjective imaging-based evaluation and be used to inform clinical decision-making (9,10). There are several major limitations of Radiomics, including lack of standardization, prospective studies and histologic validation (11). However, radiomics still will be seen as an imaging tool for accurate assessment of lesions. In CRC, radiomics can detect the tumor, predict tumor stage (12), microsatellite stability status (13) and lymph node metastasis (14,15), evaluate the response of treatment (16), and predict prognosis (17) and disease-free survival (18). Research has been conducted on the evaluation of immunohistochemical markers of cancer. One study found that multiparametric MRI radiomics signatures could noninvasively assess Ki-67 expression in rectal cancer (19), with the area under the curve (AUC) value of 0.699 [95% confidence interval (CI): 0.611–0.786]. It is necessary to further improve the performance of radiomics models. Liu *et al.* (20) showed that radiomics approaches could predict the expression level of Ki-67 in breast cancer. However, they extracted radiomics features only from a

Highlight box

Key findings

- Radiomics analysis based on preoperative magnetic resonance imaging can evaluate postoperative Ki-67, p53, and epidermal growth factor receptor (EGFR) in rectal cancer.

What is known and what is new?

- Radiomics can reveal the potential relationship between medical images and microscopic pathophysiological features.
- Rad-score based on preoperative magnetic resonance imaging can provide good evaluation of Ki-67, p53, and EGFR expression in rectal cancer.

What is the implication, and what should change now?

- The radiomics analysis employed in this study exhibited superior performance compared to that reported in other studies. Our findings contribute to the clinical decision-making related to the treatment of patients with rectal cancer.

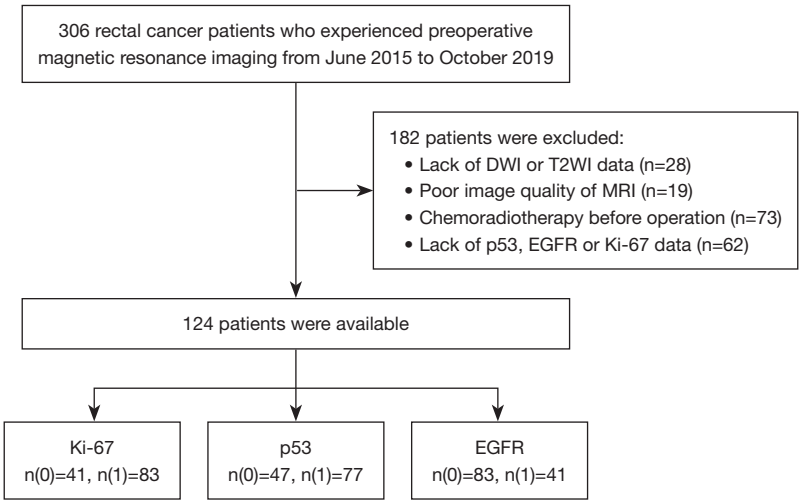

Figure 1 The workflow of the research. DWI, diffusion weighted imaging; T2WI, T2-weighted imaging; MRI, magnetic resonance imaging; EGFR, epidermal growth factor receptor; 0, negative; 1, positive.

single image of the largest cross-section of the tumor, which may result in the missing of crucial information about tumor heterogeneity. Therefore, the region of interest (ROI) should include each slice of the tumor and then be merged into the volume of interest (VOI).

Therefore, in our study, we built and validated MRI-based radiomics signatures for the noninvasive evaluation of the Ki-67, p53, and EGFR expression of rectal cancer. We present this article in accordance with the TRIPOD reporting checklist (available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-220/rc).

Methods

Patients

This retrospective study was approved by the Ethics Committee of Longyan First Affiliated Hospital of Fujian Medical University (ethical approval No. [2020] No. 036), and no patient informed consent was required. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) (21).

Data from 306 patients who underwent radical resection for rectal cancer from June 2015 to October 2019 were retrospectively collected. The inclusion criteria for patients were as follows: (I) both DWI and T2WI data available; (II) presence of primary rectal adenocarcinoma; (III) no treatment received before operation; (IV) performance of total mesorectal excision (TME) surgery; and (V) complete data on p53, EGFR, and Ki-67 expression. Meanwhile, the exclusion criteria were as follows: (I) lack of DWI or T2WI data; (II) poor image quality of MRI; (III) chemoradiotherapy before operation; and (IV) lack of p53, EGFR, or Ki-67 data. Finally, 124 patients were enrolled; the workflow for this study is displayed in *Figure 1*. Pathological TNM staging was applied in the study.

MRI data acquisition retrieval procedure

Within 1–2 weeks prior to the surgery, all patients underwent an MRI with 1.5-T MR scanners (Signa HDxt, GE HealthCare, Chicago, IL, USA) equipped with 8-channel phased-array body coils. The patients were placed in the supine position, and no bowel preparation was required before the examination. The standard rectal MRI protocol was followed and shown in *Table 1*. Axial DWI images were acquired by employing single-shot echo-planar imaging with two b-values (0 and 800 s/mm^2), with echo time (TE) =77 ms, repetition time (TR) =5,000 ms, field of view (FOV) =280 mm × 280 mm, gap =1.0 mm, matrix =128×128, and thickness =5.0 mm. Fast recovery fast spin-echo was used to acquire T2WI images with FOV =280 mm × 280 mm, TR =5,300 ms, echo train length =16, TE =116 ms, gap =0 mm, matrix =384×384, and thickness =3 mm. For additional image processing, all MRI data were extracted from the picture archiving and communication system (PACS).

Table 1 MRI protocols of T2WI and DWI

Variables	TE (ms)	TR (ms)	FOV (mm)	Thickness (mm)	Gap (mm)	Matrix
DWI	77	5,000	280×280	5.0	1.0	128×128
T2WI	116	5,300	280×280	3.0	0	384×384

MRI, magnetic resonance imaging; T2WI, T2-weighted imaging; DWI, diffusion weighted imaging; TE, echo time; TR, repetition time; FOV, field of view.

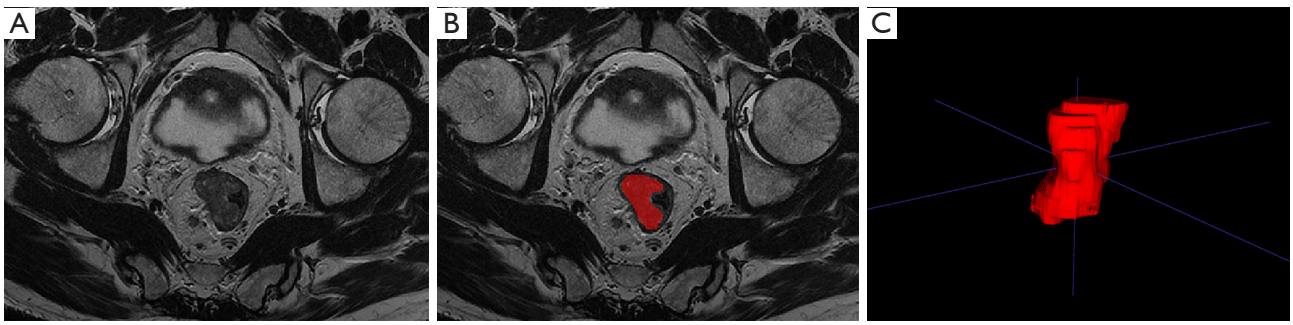

Figure 2 A 61-year-old male with T3N2M0 rectal cancer. (A) Original T2-weighted magnetic resonance image. (B) Region of interest in one slice. (C) Three-dimensional region of interest composition.

Immunohistochemical biomarkers tests

EGFR, Ki-67, and p53 expression was detected via immunohistochemistry. The fraction of positive cells in five high-power fields was used to assess the expression. A proportion of positive tumor cell nuclei ≤25% was denoted as "+", a proportion of >25–<50% as "++", and a proportion ≥50% as "+++". Moreover, in line with other studies (22,23), the cutoff threshold was 50% ++ (24-27), according to which all patients were divided into two classes (0 = negative, <50%, and 1 = positive, ≥50%).

ROI delineation and radiomics feature extraction

From the T2WI and DWI data, the ROI was drawn on each slice manually using ITK-SNAP software (www.itksnap.org) (28). ROIs were drawn along the tumor's outline and placed on the marginally high signal area on T2WI. The same tumor outline, including any signal intensity region areas of the lesion on DWI (b-value of 800 s/mm^2), was considered to be the ROI. Subsequently, the ROIs of each slice were merged automatically into the VOIs. Two radiologists, each with 10 years' experience in diagnosis of rectal cancer, worked together to complete the ROI delineation. One radiologist determined the contours and plotted the shape along the boundaries by layers. The other radiologist calibrated and verified the results before the entire tumor scope was plotted. The intraclass correlation coefficients (ICCs) of the observers were computed. The ROI delineation is shown in *Figure 2*.

Artificial Intelligence Kit software (GE HealthCare) was used to extract imaging characteristics from normalized pretreatment DWI and T2WI data with segmented VOIs, with 398 radiomics parameters being extracted from T2WI and DWI, respectively.

Feature selection method and model construction

All patients were randomly assigned to validation and training cohorts in a 3:7 ratio. To minimize discrepancies in the value scales of features, feature normalization was completed using Z-score (each feature was subtracted from the mean value and divided by the standard deviation) in the training cohorts. The multicollinearity threshold was 0.9, and the outlier removing threshold was 0.05. Abnormal values were displaced by the median value of the feature. In order to identify redundant features, two dimension reductions were performed as follows: the least absolute shrinkage and selection operator (LASSO) and the minimum redundancy maximum relevance (*mRMR*) were used to select the feature. mRMR was initially carried out for removing the unnecessary and redundant features. Following this, LASSO

was used to select the best subset of characteristics from which to build the final model. The best predictive subset of characteristics was chosen once the number of features was established, and the associated coefficients were evaluated. The radiomics score (Rad-score) signature was calculated by adding the selected features weighted by their coefficients in the validation and training cohorts. We used receiver operating characteristic (ROC) analysis to assess the model's performance and decision curve analysis (DCA) to evaluate the model's clinical usefulness.

Statistical analysis

The SPSS version 22.0 (IBM Corp., Armonk, NY, USA) was used, and the χ^2 or Fisher exact test was applied to analyze the clinical characteristics of patients. R software version 3.3.2 (The R Foundation for Statistical Computing; https://www.r-project.org) was used to carry out radiomics data analysis. The "pROC" package in R was used to analyze the ROC curve. The AUC was used to evaluate the model's performance. Accuracy, sensitivity, NPV, specificity, and positive predictive value (PPV) were identified using the Youden index to compare differences in biological characteristics. Variables with a P value <0.05 were considered statistically significant.

Results

Patient demographic data

This study included 124 patients who were categorized into training (n=86) and test (n=38) groups. The demographic parameters of the patients are listed in *Table 2*. There were 78 men (62.9%) and 46 women (37.1%), and 62.9% of the participants were over 60 years old. The majority of cancers were in the middle of rectum (n=72, 58.1%), T3 stage (n=71, 57.3%), N0 stage (n=67, 54.0%), and M0 stage (n=112, 90.3%) and with a moderate degree of differentiation (n=102, 82.3%). However, there were only significant differences in sex between the Ki-67-positive and Ki-67-negative patients (P<0.05) and no significant differences in age, location, differentiation degree, or TNM stage between the Ki-67-positive and Ki-67-negative patients, p53-positive and p53-negative patients, or between the EGFR-positive and EGFR-negative patients (P>0.05).

Feature selection and radiomics signature construction

The 30 features were chosen using LASSO and mRMR from a total of 796 features. LASSO was performed to select the best subset of characteristics to build the final model. The LASSO procedure entailed selecting a regularization parameter λ and identifying the number of the features for Ki-67, p53, and EGFR, as shown in Figure S1. The most predictive features were selected, and the coefficients associated with each feature were assessed (*Figure 3*).

Development and validation of the radiomics signature

Rad-scores of Ki-67, p53, and EGFR were calculated by adding the weights of the selected features' values by their coefficients. The Rad-scores' final formula is provided in Appendix 1. The Rad-scores from class 0 and class 1 (0 = negative, 1 = positive) were compared in the training and test groups, respectively, using the Wilcoxon test, with all P values being less than 0.05 (*Figure 4*). The performance of the three models was assessed using ROC analysis (*Figure 5*). In addition, we calculated the sensitivity, NPV, accuracy, specificity, PPV based on the Youden Index to verify that the outcomes were reliable. The detailed information on the performance of the three models is presented in *Table 3*. The Rad-score for evaluating the expression status of Ki-67 yielded AUC values of 0.91 (95% CI: 0.87–0.95) and 0.81 (95% CI: 0.66–0.96), accompanied by accuracy, sensitivity, specificity, PPV, and NPV values of 0.85, 0.80, 0.92, 0.93, and 0.78 in the training group, and of 0.75, 0.75, 0.75, 0.86, and 0.60 the test group, respectively. The evaluation of p53 expression status yielded AUC values of 0.82 (95% CI: 0.77–0.88) and 0.80 (95% CI: 0.65–0.96), along with corresponding accuracy, sensitivity, specificity, PPV, and NPV values of 0.76, 0.80, 0.72, 0.80, and 0.72 in the training group, respectively, and of 0.75, 0.78, 0.69, 0.82, and 0.64 in the test group, respectively. For EGFR expression status in both training and test groups respectively, the AUC values were 0.86 (95% CI: 0.81–0.91) and 0.76 (95% CI: 0.58–0.93), respectively. The corresponding accuracy, sensitivity, specificity, PPV, and NPV were 0.82, 0.76, 0.88, 0.82, and 0.83 for the training group, respectively, and 0.69, 0.75, 0.67, 0.53, and 0.84 for the test group, respectively.

Clinical usefulness

Figure 6 shows the DCA findings for the radiomics model. In terms of clinical applicability, the DCA outcomes were promising. The threshold of achieving prediction was 0% and 100%. The Rad-scores for predicting the expression

Table 2 Clinicopathological features of 124 patients with rectal cancer

Variables	N (%)	Ki-67, n			p53, n			EGFR, n		
		0 (n=41)	1 (n=83)	P	0 (n=47)	1 (n=77)	P	0 (n=83)	1 (n=41)	P
Gender				0.04			0.87			0.76
Male	78 (62.9)	31	47		30	48		53	25	
Female	46 (37.1)	10	36		17	29		30	16	
Age (years)				0.93			0.55			0.38
≤60	46 (37.1)	15	31		19	27		33	13	
>60	78 (62.9)	26	52		28	50		50	28	
Major location				0.54			0.10			0.16
Upper	27 (21.8)	10	17		8	19		14	13	
Middle	72 (58.1)	21	51		25	47		52	20	
Lower	25 (20.1)	10	15		14	11		17	8	
Differentiation degree				0.44			0.39			0.79
Poor	17 (13.7)	8	9		9	8		12	5	
Moderate	102 (82.3)	32	70		36	66		67	35	
Well	5 (4.0)	1	4		2	3		4	1	
pT stage				0.17			0.82			0.84
≤T2	34 (27.4)	7	27		13	21		24	10	
T3	71 (57.3)	26	45		28	43		47	24	
T4	19 (15.3)	8	11		6	13		12	7	
pN stage				0.41			0.55			0.23
N0	67 (54.0)	20	47		27	40		48	19	
N1–3	57 (46.0)	21	36		20	37		35	22	
pM stage				0.10			0.22			0.76
M0	112 (90.3)	34	78		40	72		74	38	
M1	12 (9.7)	7	5		7	5		9	3	

EGFR, epidermal growth factor receptor; pT, pathological tumor stage; pN, pathological nodal stage; pM, pathological metastasis stage; 0, negative; 1, positive.

level indicated a higher benefit in contrast to an all-or-none treatment prediction.

Discussion

In this study, we develop and validate Rad-score models consisting of preoperative T2WI and DWI data for the evaluation of the expression levels of Ki-67, p53, and EGFR of rectal cancer. The Rad-score models showed high AUC values (0.91, 0.82, 0.86) in training groups, which demonstrated excellent performance in effectively predicting the status of Ki-67, p53, and EGFR expression which may aid in informing the clinical decision-making related to rectal cancer treatment.

Rad-score models (27) have been proven to be an excellent approach for predicting Ki-67 expression in rectal cancer in recent studies, while mRMR and LASSO (29,30) are feature-selection methods that support the good performance of radiomics models. For the texture analysis of the Ki-67, p53, or EGFR expression status in

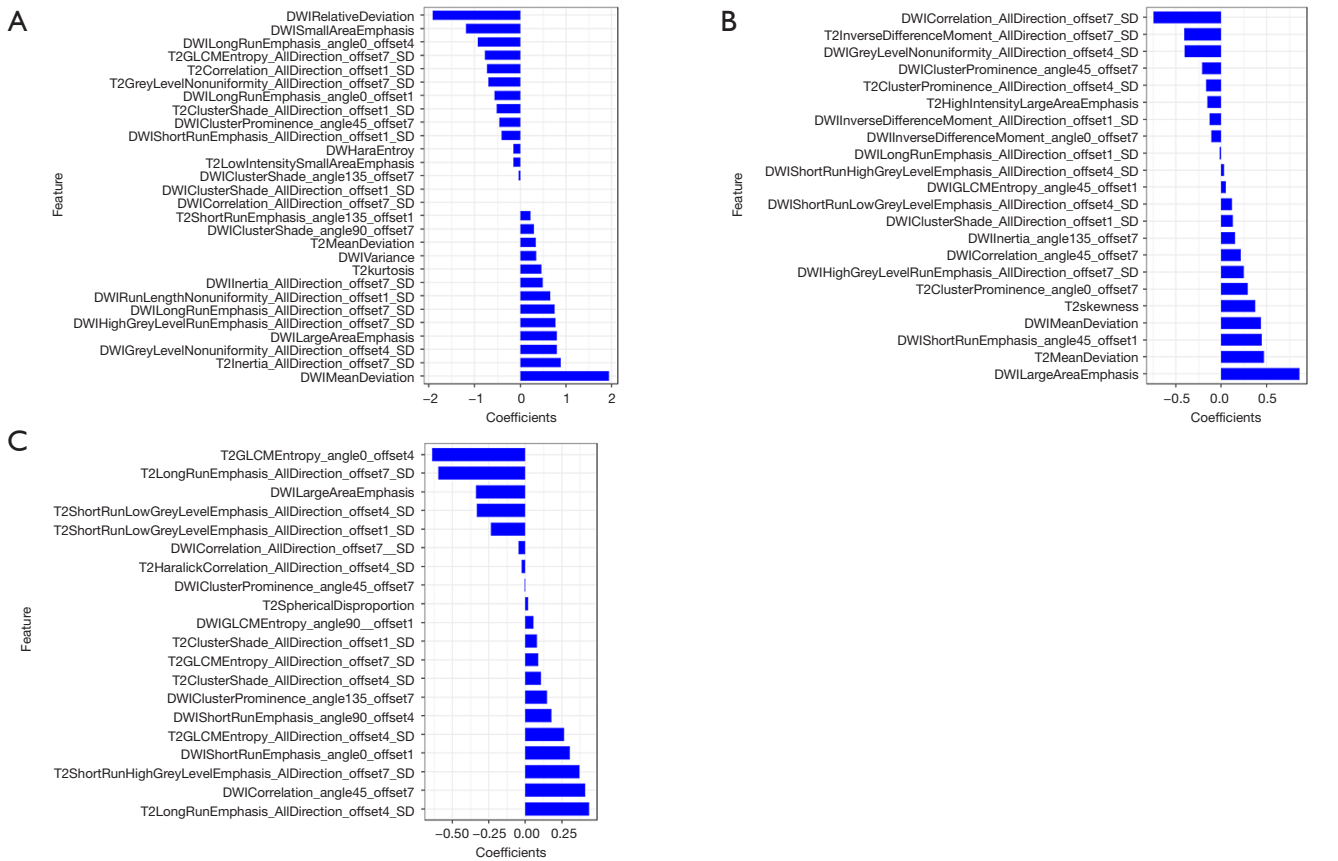

Figure 3 The subset of features with the best predictive ability and their corresponding coefficients. (A) Ki-67. (B) p53. (C) EGFR. Diffusion-weighted imaging texture parameters and T2-weighted texture parameters were features extracted from diffusion-weighted imaging and T2-weighted imaging, respectively. EGFR, epidermal growth factor receptor.

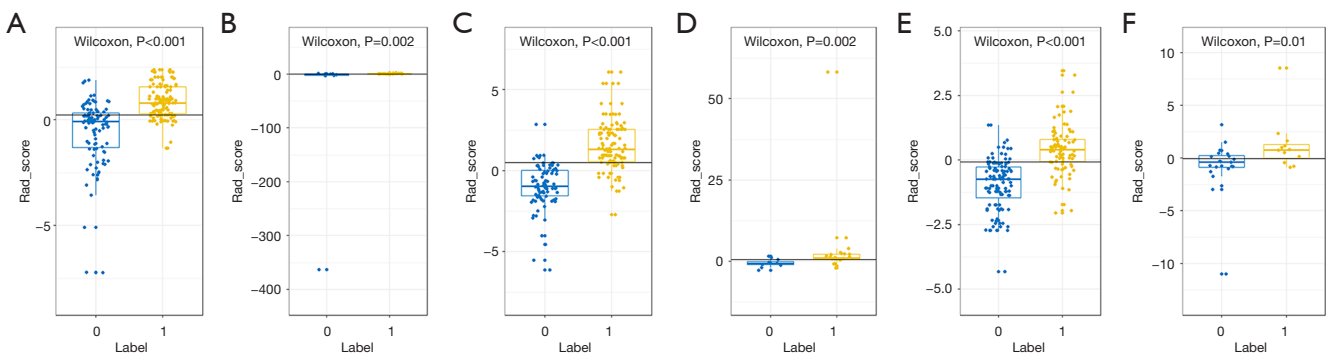

Figure 4 Wilcoxon test for comparing the radiomics score of each patient from class 0 (negative) and class 1 (positive) in the training group and testing group, respectively. (A) Ki-67 yielded P<0.001 and (B) P=0.002 in the training and testing group, respectively. (C) p53 yielded P<0.001 and (D) P=0.002 in the training and testing group, respectively. (E) EGFR yielded P<0.001 and (F) P=0.01 in training and testing group, respectively. EGFR, epidermal growth factor receptor.

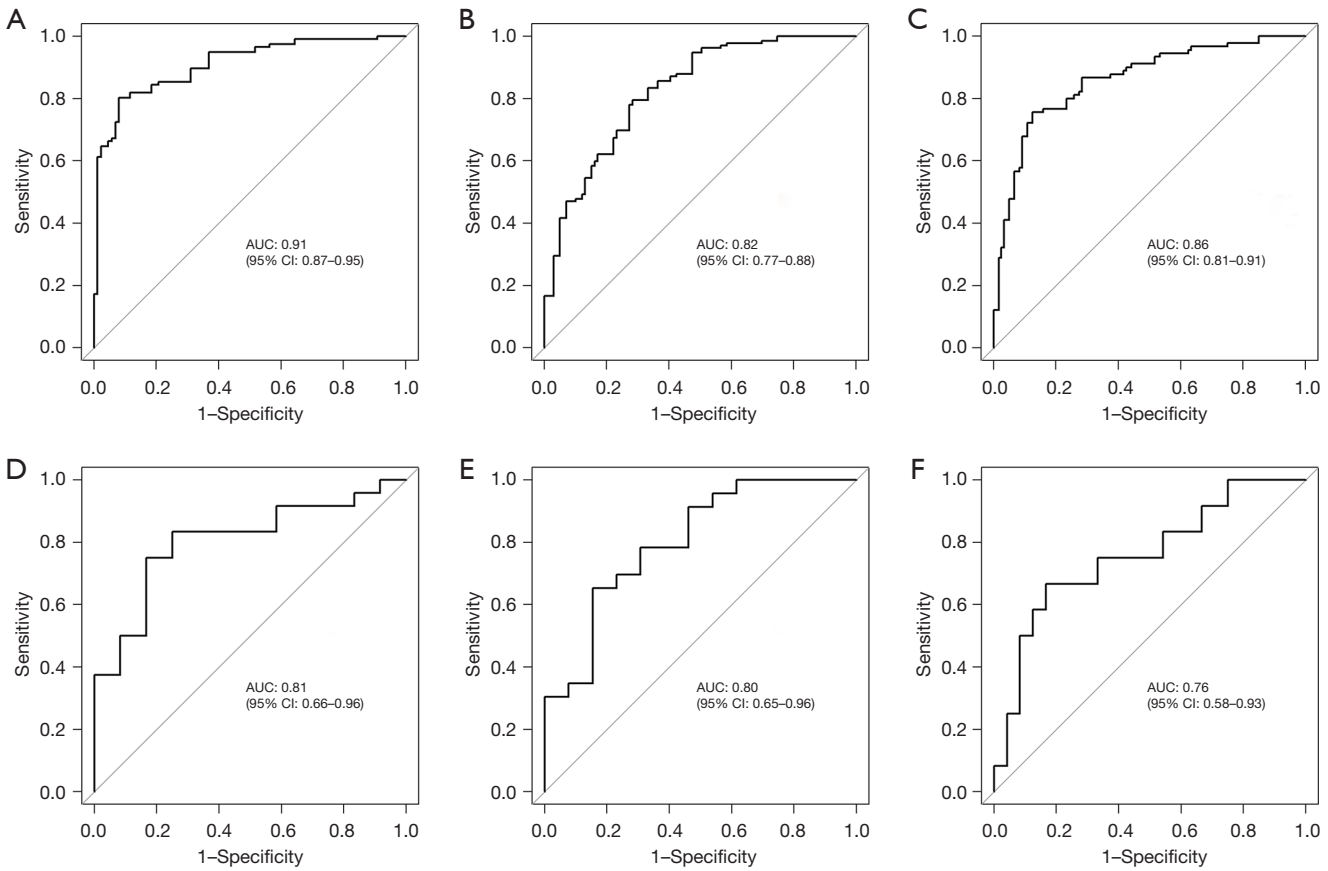

Figure 5 Receiver operating characteristic of the model's performance in the testing and training groups. (A) Ki-67, (B) p53, and (C) EGFR in the training group. (D) Ki-67, (E) p53, and (F) EGFR in the test group. EGFR, epidermal growth factor receptor; AUC, area under the curve; CI, confidence interval.

Table 3 Accuracy, sensitivity, specificity, PPV, NPV of 3 models

Variables	Accuracy	Sensitivity	Specificity	PPV	NPV
Ki-67					
Training	0.85	0.80	0.92	0.93	0.78
Test	0.75	0.75	0.75	0.86	0.60
p53					
Training	0.76	0.80	0.72	0.80	0.72
Test	0.75	0.78	0.69	0.82	0.64
EGFR					
Training	0.82	0.76	0.88	0.82	0.83
Test	0.69	0.75	0.67	0.53	0.84

PPV, positive predict value; NPV, negative predict value; EGFR, epidermal growth factor receptor.

rectal cancer, Meyer *et al.* (23) found that the maximum apparent diffusion coefficient (ADC) correlated with EGFR expression, while kurtosis and skewness correlated inversely with p53 expression; moreover, they reported there to be a close association of Ki-67 expression with ADC median, p10 and p25. However, their study used histogram features (first-order statistics features) analysis, and there were no radiomics models applied. In our study, the correlation between individual features and the level of Ki-67, p53, and EGFR was not examined directly, but Rad-score models were developed with different orders of statistics features. The first-order features were used for characterizing the distribution of gray values within the image and for the identification and classification of lesions. The second-order statistics involved the correlation of two pixels. The

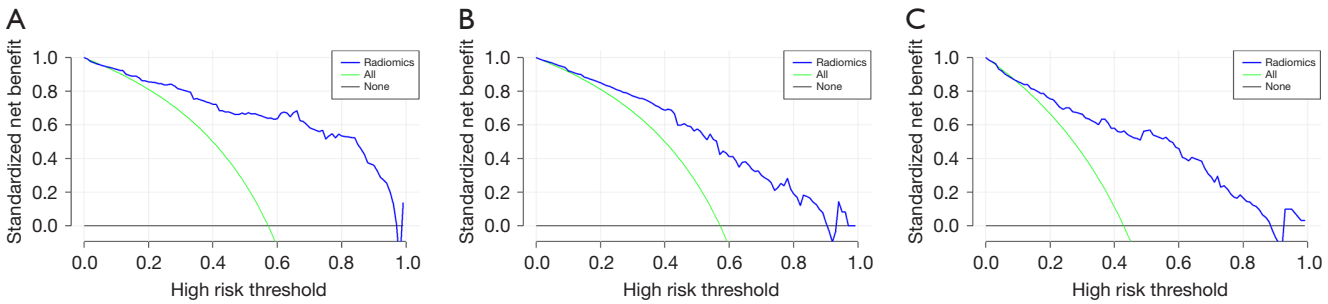

Figure 6 Decision curves for assessing the clinical usefulness of the model in the training dataset. (A) Ki-67, (B) p53, and (C) EGFR. The horizontal line represents the assumption of no positive patients. The blue curve represents the standardized net benefit and was classified as positive (high risk) by the radiomics model at each threshold probability. The green curve represents the standardized net benefit of true positives at each threshold probability. EGFR, epidermal growth factor receptor.

high-dimensional features involved the correlation of more pixels that could reflect the spatial characteristics of the texture and the heterogeneity of tumor characteristics, thus providing more detailed objective information of the cancer that correlated with tumor proliferation and prognosis. Among the optimal radiomics signatures in the model for Ki-67, the number of second-order statistics and high-dimensional features was the highest (18/28, 64.3%), followed by that of first-order features (10/28, 35.7%). In the model for p53, the second-order statistics and high-dimensional features accounted for the majority of features (19/22, 86.4%), followed first-order features (3/22, 13.6%). Finally, in the model for EGFR, all the features were second-order statistics or high-dimensional features. With the combination of fewer first-order features and a large number of second-order statistics or high-dimensional features, the radiomics models not only represented the morphological characteristics but also indicated the heterogeneities of tumors. This approach may improve performance of radiomics models. Other studies (27,31) found that "wavelet" features, a kind of high-dimensional feature, were major components of the radiomics model, which is in line with our findings. However, the AUC of 0.91 in our study indicated a better predictive ability for evaluating the expression level of Ki-67 compared to that reported in previous studies with the AUC of 0.83 on rectal cancer (27). Moreover, T2WI and DWI can provide complementary information about tumor and tissue characteristics, the Rad-scores with the combination of conventional T2WI and functional imaging DWI features may offer more useful information. In a recent study by Liu *et al.* (32), a radiomics model with high-dimensional features of T2WI and DWI data could predict pathological complete response (pCR) of rectal cancer with high accuracy. Additionally, Ma *et al.* (33) reported that a T2WI-based radiomics model could predict the N stage, degree of differentiation, and T stage of rectal cancer. Yao *et al.* (34) demonstrated that a radiomics model with features of T2WI, DWI, and another sequence could predict recurrence and metastasis in rectal cancer. Similarly, in our research, we combined T2WI and DWI features to develop and validate Rad-score models and obtained the highest AUC value and accuracy of 0.91 and 0.85. All of these studies, including our own, confirm that radiomics models may support clinical decision-making, providing noninvasive and accurate evaluation.

Radiomics provides a novel non-invasive strategy for assessing cell proliferation and has been shown to facilitate the precise assessment of tumor heterogeneity, which has an important relationship with the prognosis of various malignant tumors (35). Besides in rectal cancer, other diseases have been examined in relation to radiomics and immunohistochemical expression. Ni *et al.* (36) reported that logistic regression achieved the AUC of 0.912, and linear support vector machine (SVM) obtained the accuracy with a score of 0.884 to predict Ki-i67 expression in glioma. Similarly, in our study, Rad-scores of Ki-67 showed high AUC value of 0.91, and accuracy of 0.85 in rectal cancer. Moon *et al.* (37) developed clinical-radiomic machine learning models with the AUC and accuracy of 0.820 and 0.867 for predicting Ki-67, and 0.858 and 0.857 to predict p53 expression in meningioma using MRI features. The radiomics models in the study of Wu *et al.* (38) yielded AUC values of 0.74 and 0.95 in the test and training set,

respectively, in breast ductal carcinoma *in situ*. In our study, the Rad-scores of p53 produced AUC values of 0.82 and 0.80 in the training and test set, indicating the good predictive ability of the expression of p53 in rectal cancer. Moreover, the radiomics model of Wu *et al.* (39) yielded an AUC of 0.950 and could predict the expression of EGFR of peripheral lung cancer. In our study, Rad-scores for evaluating EGFR expression status in both training and test groups, the AUC values were 0.86 and 0.76, respectively. Furthermore, we used DCA to quantify net benefits of Rad-cores under different threshold probabilities for predicting Ki-67, p53, EGFR expression, and revealed that the net benefits of Rad-cores were all higher than "None" and "All". Radiomics models with optimal radiomics signatures have shown the ability to predict the status of immunohistochemical markers in various malignant, and may thus be used to assist in clinical decision-making in the different tumors' treatment.

Radiomics, as a complementary tool for clinicians, has been widely studied in the management of cancer patients (40). Radiomics is a more accurate prediction of tumor cells and Ki-67, p53 and EGFR expression, which may help prognosticate the heterogeneous clinical behavior, and reduce time-consuming and costly procedure, particularly for patients who may be able to avoid or benefit from aggressive treatment options.

This study involved certain limitations which should be acknowledged. Firstly, our study had a relatively limited sample size and 124 patients were recruited from a single center, and a much larger patient sample and a dataset from different centers need to be employed in the future. Secondly, the MRI data were extracted from a single MR scanner, and the use of multiple MR scanners would produce more persuasive results. Additionally, we only employed radiomics signature to predict the expression levels of Ki-67, p53, and EGFR of tumor and did not consider the predictive value of other clinical indicators, and clinical indicators should be added to further investigation in the future.

Conclusions

Radiomics signatures for evaluating postoperative immunohistochemical markers of rectal cancer based on preoperative MRI were established and validated in this study. Overall, radiomics has the potential to noninvasively predict the Ki-67, p53, and EGFR of rectal cancer before operation is performed.

Acknowledgments

Funding: This study was supported by the Fujian Provincial Natural Science Foundation (No. 2020J011326); and the Key Clinical Specialty Discipline Construction Program of Fujian (No. Fujian Medical Policy Letter 2018-145).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-220/rc

Data Sharing Statement: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-220/dss

Peer Review File: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-220/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-220/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Ethics Committee of The Longyan First Affiliated Hospital of Fujian Medical University (ethical approval No. [2020] No. 036). The requirement for informed consent was waived due to the retrospective nature of the study.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics

2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209-49.
2. Terzi C, Canda AE, Sagol O, et al. Survivin, p53, and Ki-67 as predictors of histopathologic response in locally advanced rectal cancer treated with preoperative chemoradiotherapy. Int J Colorectal Dis 2008;23:37-45.
3. Kasprzak A. Prognostic Biomarkers of Cell Proliferation in Colorectal Cancer (CRC): From Immunohistochemistry to Molecular Biology Techniques. Cancers (Basel) 2023;15:4570.
4. Wang H, Chen Q, Liu Q, et al. Master regulator: p53's pivotal role in steering NK-cell tumor patrol. Front Immunol 2024;15:1428653.
5. Minciuna CE, Tanase M, Manuc TE, et al. The seen and the unseen: Molecular classification and image based-analysis of gastrointestinal cancers. Comput Struct Biotechnol J 2022;20:5065-75.
6. Hur H, Kim NK, Min BS, et al. Can a biomarker-based scoring system predict pathologic complete response after preoperative chemoradiotherapy for rectal cancer? Dis Colon Rectum 2014;57:592-601.
7. Gürses B, Böge M, Altınmakas E, et al. Multiparametric MRI in rectal cancer. Diagn Interv Radiol 2019;25:175-82.
8. Boot J, Gomez-Munoz F, Beets-Tan RGH. Imaging of rectal cancer. Radiologe 2019;59:46-50.
9. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics. Phys Med 2017;38:122-39.
10. Bourgier C, Colinge J, Ailleres N, et al. Radiomics: Definition and clinical development. Cancer Radiother 2015;19:532-7.
11. Caruso D, Polici M, Zerunian M, et al. Radiomics in Oncology, Part 2: Thoracic, Genito-Urinary, Breast, Neurological, Hematologic and Musculoskeletal Applications. Cancers (Basel) 2021;13:2681.
12. Fan L, Wu H, Wu Y, et al. Preoperative prediction of rectal Cancer staging combining MRI deep transfer learning, radiomics features, and clinical factors: accurate differentiation from stage T2 to T3. BMC Gastroenterol 2024;24:247.
13. Bodalal Z, Hong EK, Trebeschi S, et al. Non-invasive CT radiomic biomarkers predict microsatellite stability status in colorectal cancer: a multicenter validation study. Eur Radiol Exp 2024;8:98.
14. Sun Y, Lu Z, Yang H, et al. Prediction of lateral lymph node metastasis in rectal cancer patients based on MRI using clinical, deep transfer learning, radiomic, and fusion models. Front Oncol 2024;14:1433190.
15. Zheng Y, Chen X, Zhang H, et al. Multiparametric MRI-based radiomics nomogram for the preoperative prediction of lymph node metastasis in rectal cancer: A two-center study. Eur J Radiol 2024;178:111591.
16. Chen L, Zhu W, Zhang W, et al. Magnetic resonance imaging radiomics-based prediction of severe inflammatory response in locally advanced rectal cancer patients after neoadjuvant radiochemotherapy. Langenbecks Arch Surg 2024;409:218.
17. Mao J, Ye W, Ma W, et al. Prediction by a multiparametric magnetic resonance imaging-based radiomics signature model of disease-free survival in patients with rectal cancer treated by surgery. Front Oncol 2024;14:1255438.
18. Guo X, He Y, Yuan Z, et al. Association Analysis Between Intratumoral and Peritumoral MRI Radiomics Features and Overall Survival of Neoadjuvant Therapy in Rectal Cancer. J Magn Reson Imaging 2024. [Epub ahead of print]. doi: 10.1002/jmri.29396.
19. Meng X, Xia W, Xie P, et al. Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer. Eur Radiol 2019;29:3200-9.
20. Liu J, Yan C, Liu C, et al. Predicting Ki-67 expression levels in breast cancer using radiomics-based approaches on digital breast tomosynthesis and ultrasound. Front Oncol 2024;14:1403522.
21. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310:2191-4.
22. Liu S, Shi H, Ji C, et al. CT textural analysis of gastric cancer: correlations with immunohistochemical biomarkers. Sci Rep 2018;8:11844.
23. Meyer HJ, Höhn A, Surov A. Histogram analysis of ADC in rectal cancer: associations with different histopathological findings including expression of EGFR, Hif1-alpha, VEGF, p53, PD1, and KI 67. A preliminary study. Oncotarget 2018;9:18510-7.
24. Yang Y, Li J, Jin L, et al. Independent Correlation Between Ki67 Index and Circulating Tumor Cells in the Diagnosis of Colorectal Cancer. Anticancer Res 2017;37:4693-700.
25. Tong G, Zhang G, Liu J, et al. Cutoff of 25% for Ki67 expression is a good classification tool for prognosis in colorectal cancer in the AJCC-8 stratification. Oncol Rep 2020;43:1187-98.
26. Li P, Xiao ZT, Braciak TA, et al. Association between Ki67 Index and Clinicopathological Features in Colorectal Cancer. Oncol Res Treat 2016;39:696-702.

27. Yao X, Ao W, Zhu X, et al. A novel radiomics based on multi-parametric magnetic resonance imaging for predicting Ki-67 expression in rectal cancer: a multicenter study. BMC Med Imaging 2023;23:168.
28. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31:1116-28.
29. Abbaspour S, Barahman M, Abdollahi H, et al. Multimodality radiomics prediction of radiotherapy-induced the early proctitis and cystitis in rectal cancer patients: a machine learning study. Biomed Phys Eng Express 2023.
30. Li Z, Huang H, Zhao Z, et al. Development and Validation of a Nomogram Based on DCE-MRI Radiomics for Predicting Hypoxia-Inducible Factor 1α Expression in Locally Advanced Rectal Cancer. Acad Radiol 2024;S1076-6332(24)00300-3.
31. Wang H, Chen X, Ding J, et al. Novel multiparametric MRI-based radiomics in preoperative prediction of perirectal fat invasion in rectal cancer. Abdom Radiol (NY) 2023;48:471-85.
32. Liu Z, Zhang XY, Shi YJ, 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.
33. Ma X, Shen F, Jia Y, et al. MRI-based radiomics of rectal cancer: preoperative assessment of the pathological features. BMC Med Imaging 2019;19:86.
34. Yao X, Zhu X, Deng S, et al. MRI-based radiomics for preoperative prediction of recurrence and metastasis in rectal cancer. Abdom Radiol (NY) 2024;49:1306-19.
35. Li B, Zhu J, Wang Y, et al. Radiomics nomogram based on CT radiomics features and clinical factors for prediction of Ki-67 expression and prognosis in clear cell renal cell carcinoma: a two-center study. Cancer Imaging 2024;24:103.
36. Ni J, Zhang H, Yang Q, et al. Machine-Learning and Radiomics-Based Preoperative Prediction of Ki-67 Expression in Glioma Using MRI Data. Acad Radiol 2024;31:3397-405.
37. Moon CM, Lee YY, Kim DY, et al. Preoperative prediction of Ki-67 and p53 status in meningioma using a multiparametric MRI-based clinical-radiomic model. Front Oncol 2023;13:1138069.
38. Wu L, Zhao Y, Lin P, et al. Preoperative ultrasound radiomics analysis for expression of multiple molecular biomarkers in mass type of breast ductal carcinoma in situ. BMC Med Imaging 2021;21:84.
39. Wu L, Li J, Ruan X, et al. Prediction of VEGF and EGFR Expression in Peripheral Lung Cancer Based on the Radiomics Model of Spectral CT Enhanced Images. Int J Gen Med 2022;15:6725-38.
40. Caruso D, Polici M, Zerunian M, et al. Radiomics in Oncology, Part 1: Technical Principles and Gastrointestinal Application in CT and MRI. Cancers (Basel) 2021;13:2522.

Cite this article as: Li Q, Liu J, Li W, Qiu M, Zhuo X, You Q, Qiu S, Lin Q, Liu Y. Magnetic resonance imaging-based radiomics in predicting the expression of Ki-67, p53, and epidermal growth factor receptor in rectal cancer. J Gastrointest Oncol 2024;15(5):2088-2099. doi: 10.21037/jgo-24-220

Appendix 1 Formulas of radscore

Ki-67

Radscore = -1.185 * DWISmallAreaEmphasis + 1.938 * DWIMeanDeviation + 0.803 * DWIGreyLevelNonuniformity_AllDirection_offset4_SD + 0.348 * DWIVariance + 0.798 * DWILargeAreaEmphasis + 0.651 * DWIRunLengthNonuniformity_AllDirection_offset1_SD + -0.727 * T2Correlation_AllDirection_offset1_SD + 0.226 * T2ShortRunEmphasis_angle135_offset1 + 0.884 * T2Inertia_AllDirection_offset7_SD + -0.773 * T2GLCMEntropy_AllDirection_offset7_SD + 0 * DWICorrelation_AllDirection_offset7_SD + -0.694 * T2GreyLevelNonuniformity_AllDirection_offset7_SD + -0.925 * DWILongRunEmphasis_angle0_offset4 + -0.511 * T2ClusterShade_AllDirection_offset1_SD + -0.039 * DWIClusterShade_angle135_offset7 + -0.453 * DWIClusterProminence_angle45_offset7 + 0.748 * DWILongRunEmphasis_AllDirection_offset7_SD + -0.146 * T2LowIntensitySmallAreaEmphasis + -0.002 * DWIClusterShade_AllDirection_offset1_SD + -0.148 * DWIHaraEntroy + -0.559 * DWILongRunEmphasis_angle0_offset1 + 0.298 * DWIClusterShade_angle90_offset7 + -0.408 * DWIShortRunEmphasis_AllDirection_offset1_SD + 0.497 * DWIInertia_AllDirection_offset7_SD + 0.466 * T2kurtosis + -1.919 * DWIRelativeDeviation + 0.771 * DWIHighGreyLevelRunEmphasis_AllDirection_offset7_SD + 0.335 * T2MeanDeviation + 0.593

P53

Radscore = 0.864 * DWILargeAreaEmphasis + -0.126 * DWIInverseDifferenceMoment_AllDirection_offset1_SD + -0.744 * DWICorrelation_AllDirection_offset7_SD + 0.471 * T2MeanDeviation + 0.117 * DWIShortRunLowGreyLevelEmphasis_AllDirection_offset4_SD + -0.402 * DWIGreyLevelNonuniformity_AllDirection_offset4_SD + -0.164 * T2ClusterProminence_AllDirection_offset4_SD + -0.208 * DWIClusterProminence_angle45_offset7 + 0.436 * DWIMeanDeviation + -0.11 * DWIInverseDifferenceMoment_angle0_offset7 + -0.018 * DWILongRunEmphasis_AllDirection_offset1_SD + 0.128 * DWIClusterShade_AllDirection_offset1_SD + 0.033 * DWIShortRunHighGreyLevelEmphasis_AllDirection_offset4_SD + -0.405 * T2InverseDifferenceMoment_AllDirection_offset7_SD + 0.215 * DWICorrelation_angle45_offset7 + -0.151 * T2HighIntensityLargeAreaEmphasis + 0.295 * T2ClusterProminence_angle0_offset7 + 0.445 * DWIShortRunEmphasis_angle45_offset1 + 0.252 * DWIHighGreyLevelRunEmphasis_AllDirection_offset7_SD + 0.054 * DWIGLCMEntropy_angle45_offset1 + 0.376 * T2skewness + 0.155 * DWIInertia_angle135_offset7 + 0.088

EGFR

Radscore = 0.183 * DWIClusterShade_angle45_offset7 + -0.282 * DWIClusterProminence_angle90_offset7 + 1.808 * DWIkurtosis + 1.508 * T2HaralickCorrelation_AllDirection_offset7_SD + -0.494 * DWILongRunHighGreyLevelEmphasis_AllDirection_offset4_SD + 0.618 * T2GreyLevelNonuniformity_AllDirection_offset7_SD + -1.544 * T2ClusterShade_angle0_offset7 + -0.517 * DWIHaralickCorrelation_angle45_offset7 + -0.381 * T2LongRunEmphasis_AllDirection_offset4_SD + 0.108 * T2GLCMEntropy_angle135_offset7 + -0.238 * DWIShortRunHighGreyLevelEmphasis_AllDirection_offset7_SD + 1.321 * DWISmallAreaEmphasis + -0.763 * DWIHighGreyLevelRunEmphasis_AllDirection_offset4_SD + 0.584 * DWIhistogramEnergy + -0.236 * DWICorrelation_angle0_offset7 + 0.285 * DWIShortRunEmphasis_angle90_offset7 + 0.943 * DWIClusterProminence_AllDirection_offset1_SD + 0.876 * T2MinIntensity + -0.89 * DWIInverseDifferenceMoment_AllDirection_offset1_SD + -0.243 * T2Correlation_AllDirection_offset1_SD + 1.101 * DWIHaralickCorrelation_angle135_offset7 + -0.75 * DWIMeanDeviation + 0.866 * DWIShortRunEmphasis_angle90_offset4 + -0.027 * DWILongRunEmphasis_AllDirection_offset4_SD + -0.022 * T2LargeAreaEmphasis + 0.282

https://dx.doi.org/10.21037/jgo-24-220

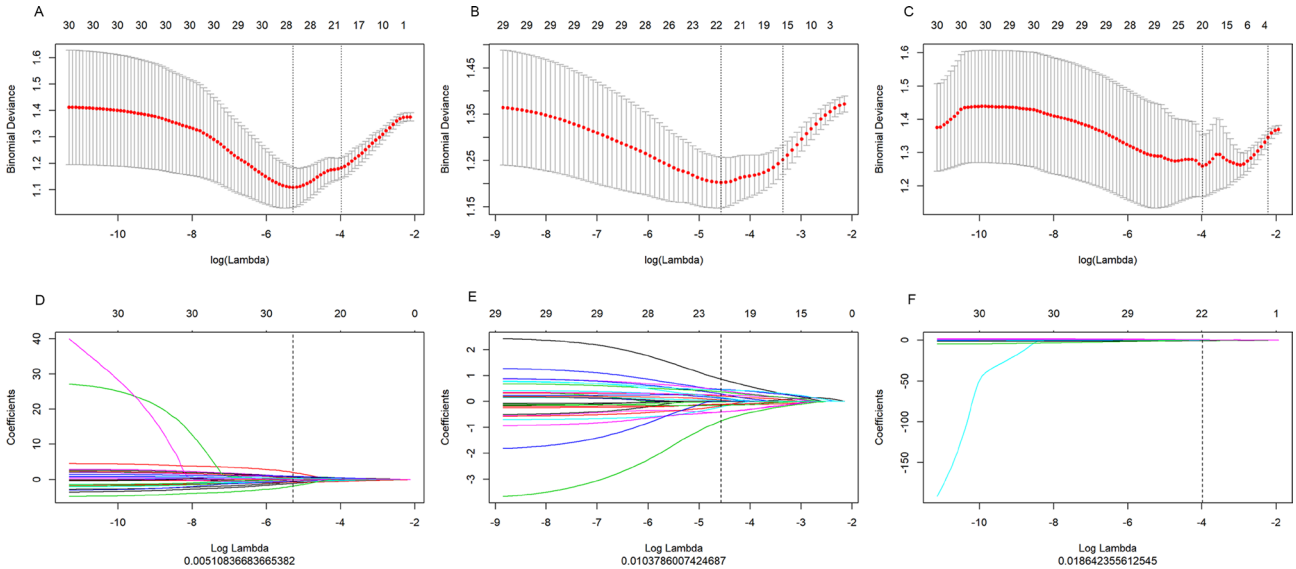

Figure S1 The LASSO procedure entailed selecting a regularization parameter λ and identifying the number of the features for Ki-67, p53, and EGFR. (A,D) Ki-67. (B,E) p53. (C,F) and EGFR.

https://dx.doi.org/10.21037/jgo-24-220