



## OPEN Development of a CT radiomics model for detection of bladder invasion by colorectal carcinoma

Jingui Wang<sup>1,3</sup>, Kexin Wang<sup>2,3</sup>, Junling Zhang<sup>1</sup>, Yingchao Wu<sup>1</sup>, Yong Jiang<sup>1</sup>, Guowei Chen<sup>1</sup>, Zhanbing Liu<sup>1</sup>, Tao Wu<sup>1</sup>, Yuanlian Wan<sup>1</sup>, Xiaoying Wang<sup>2✉</sup> & Xin Wang<sup>1✉</sup>

To investigate the feasibility of a radiomics model for the detection of bladder invasion (BI) by colorectal cancer (CRC) on CT images. Ninety-six patients with CRC and a suspicion of BI who underwent tumor resection with partial or total cystectomy were reviewed. The 96 patients were randomly assigned to the training dataset ( $n = 68$ ) or test dataset ( $n = 28$ ) at a ratio of 7:3. The CT images were reviewed by two experienced radiologists, who provided a CT impression of the invasion of the bladder by CRC. A region of interest (ROI) on the CT images for each case was manually labeled by two radiologists. A radiomics model was constructed using a Categorical Boosting (CatBoost) classifier. The predicted probability by CatBoost was used to evaluate the efficacy of the radiomics model. The areas under the curve (AUCs) of the receiver operating characteristic were compared between the radiomics model and the CT impression. In the training dataset, the AUC of the radiomic model [0.864 (95% CI: 0.778, 0.951)] was significantly greater than that of CT impression [0.678 (95% CI: 0.569, 0.786),  $P = 0.007$ ]. In the test dataset, the AUC of the radiomic model [0.883 (95% CI: 0.699, 1.000)] was also significantly greater than that of CT impression [0.570 (95% CI: 0.370, 0.770),  $P = 0.040$ ]. It is feasible to use radiomics models for the prediction of BI by CRC, which might perform better than human radiologists.

**Keywords** Colorectal cancer, Invasion, Bladder, Computed tomography, Radiomics

### Abbreviations

ACC	Accuracy
AUCs	Areas under the curve
BI	Bladder invasion
CatBoost	Categorical Boosting
CRC	Colorectal cancer
GIST	Gastrointestinal stromal tumor
ICCs	Intraclass correlation coefficients
MVR	Multivisceral resection
NLR	Negative likelihood ratio
NPV	Negative predictive value
PCA	Principal component analysis
PLR	Positive likelihood ratio
PPV	Positive predictive value
PR	Precision-recall
ROC	Receiver operating characteristic
ROI	Region of interest
SEN	Sensitivity
SPE	Specificity

Colorectal cancer (CRC) is the among the most common cancers worldwide, with at least 1.8 million new cases diagnosed annually<sup>1</sup>. Approximately 10% of these patients need to undergo multivisceral resection (MVR) because the tumor is adherent to contiguous organs or structures at the time of initial diagnosis<sup>2</sup>. However, true

<sup>1</sup>Department of Gastrointestinal Surgery, Peking University First Hospital, NO. 8 Xishiku Street, Xicheng District, Beijing 100034, China. <sup>2</sup>Department of Radiology, Peking University First Hospital, NO. 8 Xishiku Street, Xicheng District, Beijing 100034, China. <sup>3</sup>Jingui Wang and Kexin Wang contributed equally to this work. ✉email: wangxiaoying@bjmu.edu.cn; 03027@pkufh.com

malignant invasion to adjacent organs or structures was reported pathohistologically in 30–70% of patients who underwent MVR, while only inflammatory adhesion was detected in the remaining patients<sup>3</sup>. Therefore, MVR performed during curative resection for patients with CRC inflammation adhering to their surroundings is often overtreatment.

Current imaging modalities for preoperative assessment of bladder invasion (BI) face critical limitations. While contrast-enhanced CT is widely used for staging, retrospective studies demonstrate its sensitivity for macroscopic BI detection is only 64.3%, with pathological confirmation of malignancy in merely 28.6% of resected specimens<sup>4</sup>. CT cannot reliably differentiate inflammatory adhesions from true tumor infiltration, leading to overestimation of invasion. Cystoscopy, though adjunctively employed, detects mucosal abnormalities (e.g., edema) but fails to predict transmural invasion, resulting in 53% false-positive rates for malignancy<sup>4,5</sup>. These limitations underscore the unmet need for objective preoperative tools to stratify adhesion severity.

It is well acknowledged that extensive operation is often associated with a high rate of complications. Complication rates have been reported as 20–42% after MVR<sup>6</sup>. Yuji Nakafusa et al. revealed that the overall morbidity rate was 49.1% after MVR, while 17.8% after standard operation<sup>7</sup>. As the most common organ involved in MVR (53.2%)<sup>8</sup>, the bladder was partially or totally resected depending on the site and extent of bladder involvement. Compared with surgery for primary bladder cancer, urinary tract reconstruction after cystectomy is more susceptible to complications when performed via MVR for CRC<sup>9</sup>. Unfortunately, patients with inflammatory adherent CRC who undergo MVR cannot achieve oncological benefits<sup>3</sup> and suffer from a greater risk of perioperative complications and poor postoperative quality of life.

However, since imaging examination methods are limited, diagnosing the nature of adhesions preoperatively is still a challenge<sup>10</sup>. Intraoperative identification of adhesions as inflammatory or malignant is often inaccurate, and frozen sectioning is often ineffective<sup>11,12</sup>. Intraoperative attempts to release adhesions often result in transection of the tumor, which leads to higher local recurrence rates and poorer overall survival in these CRC patients than in patients undergoing en bloc MVR<sup>2,6,13–15</sup>. Thus, the preoperative evaluation of locally advanced CRC, including those with a suspicion of BI, has become increasingly important and relevant, especially in the era of minimally invasive surgery for CRC and neoadjuvant chemoradiation for rectal cancer.

The term “radiomics”, which refers to the process of converting medical images into high-dimensional data by high-throughput extraction of quantitative features and subsequent data analysis for decision-making, has recently attracted increasing interest<sup>16</sup>. Currently, radiomics has been widely applied in tumor diagnosis, prognosis assessment and treatment response prediction<sup>17</sup>. To the best of our knowledge, no studies have focused on applying radiomics models to evaluate BI in patients with CRC.

Therefore, the purpose of this study was to investigate the feasibility of a radiomics model for the detection of BI by CRC on CT images.

## Materials and methods

### Data enrollment

This retrospective study was approved by the local Institutional Review Board (Peking University First Hospital 2019–170), and all methods were performed in accordance with the relevant guidelines and regulations. Due to the retrospective nature of the study, Peking University First Hospital waived the need of obtaining informed consent. CT images of the abdomen between December 2009 and September 2021 were retrospectively collected. The inclusion criteria were as follows: (a) underwent primary CRC resection combined with partial or total cystectomy, (b) had clinical information available, and (c) had postoperative pathology to determine whether invasion of the bladder by CRC was present. The exclusion criteria were as follows: (a) CT data were obtained two weeks before surgery, (b) CT images did not fulfill the requirements of the study, (c) preoperative therapy (radiotherapy, chemotherapy or chemoradiotherapy).

A total of 96 eligible patients were ultimately included in this study; 35 had bladder invasion [BI(+)], and 61 had no bladder invasion [BI(-)]. The 96 patients were randomly assigned to the training dataset ( $n = 68$ ) or test dataset ( $n = 28$ ) at a ratio of 7:3 (Fig. 1).

### CT acquisition parameters

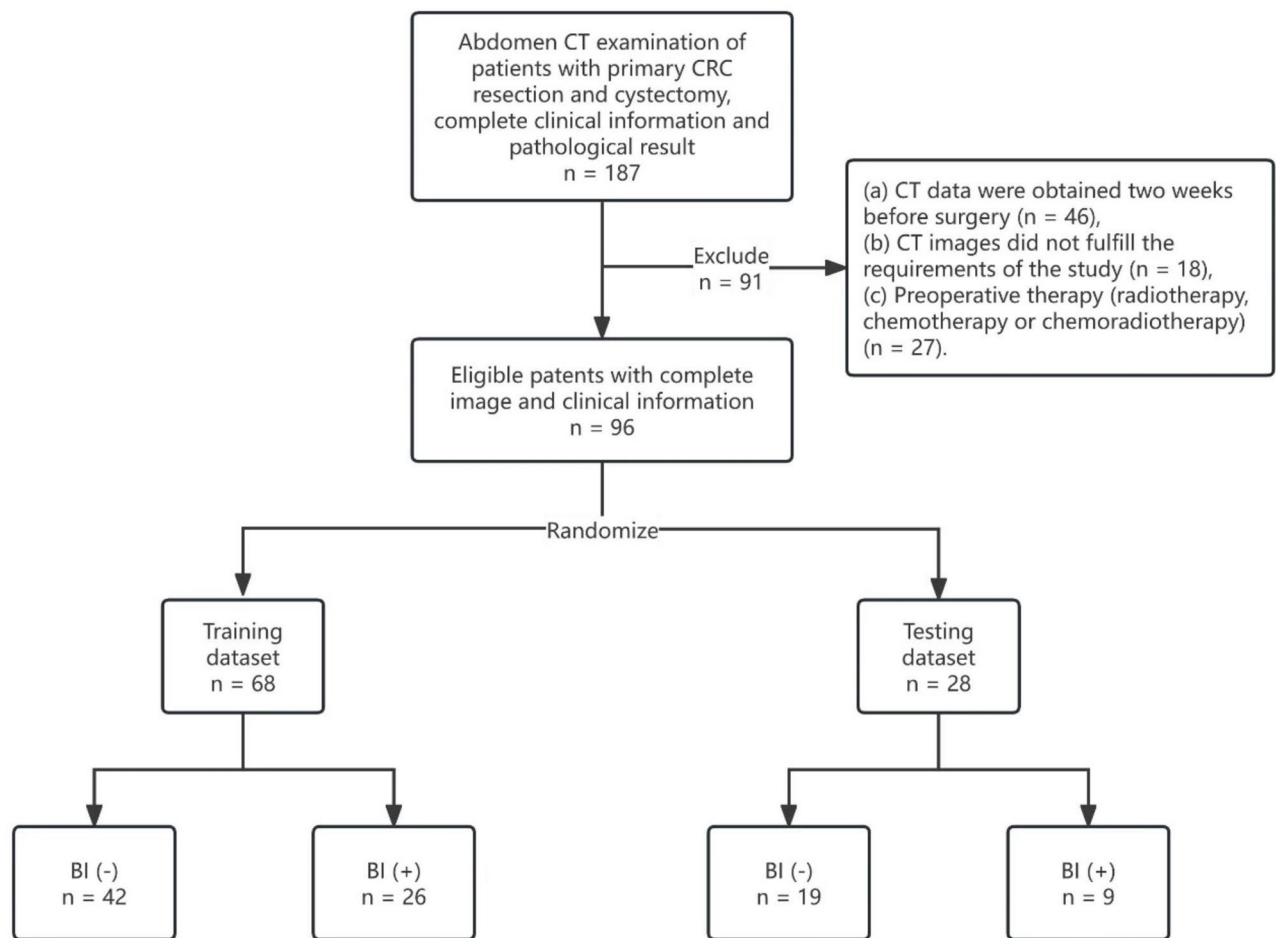
The CT images were acquired from six CT scanners. The detailed scanning parameters are shown in Table 1. There were no significant differences in the scanning parameters between the training dataset and the test dataset (all  $P > 0.05$ ).

### Clinical information

The age, sex, BMI, CA19-9, CEA, tumor location, tumor size, and differentiation of the tumors are shown in Table 2. The most common site of CRC was the sigmoid colon (62.5%), followed by the rectum (34.4%). There were no statistically significant differences in clinical information between the training and test datasets (all  $P > 0.05$ ), except for CA19-9 ( $P = 0.008$ ).

### CT impression by visual assessment of two radiologists

The CT images were reviewed by two experienced radiologists. The positive findings included CRC expansion to the bladder, thickening of the bladder wall, and stranding of the surrounding tissue. Bladder invasion was diagnosed only if direct CRC expansion to the bladder wall was observed. The presence of secondary signs (bladder wall thickening or perivesical stranding) supported the diagnosis but were not independently sufficient. The two radiologists gave the CT impression in consensus that there existed invasion of the bladder by the CRC.



**Fig. 1.** Flow chart of patient enrollment. BI(-): absence of bladder invasion. BI(+): presence of bladder invasion.

### Region of interest for radiomics model

The region of interest (ROI) on the CT images was manually labeled by two radiologists (an intern in radiology training and a radiologist with 30 years of experience) with ITK-Snap software (<http://www.itksnap.org>). A rectangle shape of the ROI was annotated at the area where the bladder and the CRC were closest (Fig. 2). Radiologists manually delineated ROIs by marking key points to form minimum bounding cuboids around tumor regions. These cuboids were not constrained to align with the image axes. ROIs were variable in size (mean 25311 mm<sup>3</sup>, range: 13054–58767 mm<sup>3</sup>) to accommodate tumor heterogeneity.

### Development of the radiomic model

All CT images were resampled to isotropic voxels (1 × 1 × 1 mm<sup>3</sup>) using B-spline interpolation to eliminate variations in spatial resolution caused by differences in slice thickness or reconstruction kernels. ROIs were converted to binary format and smoothed using morphological operations (e.g., dilation/erosion with a 3 × 3 kernel) to reduce edge artifacts. HU values were truncated to a range of -100 to 400 to exclude irrelevant tissues (e.g., air, bone). A fixed bin number (64 bins) was applied to standardize the gray-level range across all scans, ensuring comparability of texture-based features between datasets. All preprocessing adhered to the Image Biomarker Standardization Initiative (IBSI) recommendations to ensure methodological transparency and cross-study comparability. Feature extraction from these images was carried out using the PyRadiomics package in Python, with further details available in the PyRadiomics documentation. This package enabled the calculation of 14 shape-related features, 18 primary-level features, and 70 textural features. To standardize these features, the Z score normalization technique was employed. To study the reproducibility of the radiomic features, the 68 patients in the training cohort were labeled again by Reader A and Reader B.

Inter- and intraobserver reliability were assessed using intraclass correlation coefficients (ICCs), derived from a two-way random effects model. Radiomics features demonstrating excellent reliability (with ICCs greater than 0.85) were deemed robust and subsequently chosen for model construction. The selected variables were then shifted to zero centered and scaled to have unit variance. Then, principal component analysis (PCA) was performed to reduce the dimensionality of the features. Principal components (PCs) explaining 95% of the cumulative variance were retained to balance information preservation and model simplicity. The top 15 PCs

	Overall	Training dataset	Test dataset	P value
	(n = 96)	(n = 68)	(n = 28)	
Manufacture				
GE medical systems	71 (74.0%)	48 (70.6%)	23 (82.1%)	0.494
Philips	16 (16.7%)	13 (19.1%)	3 (10.7%)	
SIEMENS	9 (9.4%)	7 (10.3%)	2 (7.1%)	
Model name				
Brilliance 64	11 (11.5%)	8 (11.8%)	3 (10.7%)	0.383
Discovery CT750 HD	54 (56.3%)	37 (54.4%)	17 (60.7%)	
iCT 256	5 (5.2%)	5 (7.4%)	0 (0%)	
LightSpeed VCT	10 (10.4%)	5 (7.4%)	5 (17.9%)	
LightSpeed16	7 (7.3%)	6 (8.8%)	1 (3.6%)	
SOMATOM definition flash	9 (9.4%)	7 (10.3%)	2 (7.1%)	
Reconstruction diameter (mm)				
Median [Q1, Q3]	372 [360,387]	375 [361,389]	367 [355,385]	0.283
Slice thickness (mm)				
Median [Q1, Q3]	1.25 [1.25,1.25]	1.25 [1.00,1.25]	1.25 [1.25,1.25]	0.133
Slice spacing (mm)				
Median [Q1, Q3]	5.00 [1.00,5.00]	5.00 [1.00,5.00]	5.00 [1.00,5.00]	0.739
Pixel spacing (mm)				
Median [Q1, Q3]	0.727 [0.703,0.755]	0.732 [0.705,0.759]	0.716 [0.692,0.752]	0.319
Tube current (mA)				
Median [Q1, Q3]	300 [300,332]	300 [300,316]	300 [300,350]	0.334
Exposure time (ms)				
Median [Q1, Q3]	800 [749,800]	800 [749,800]	800 [749,800]	0.725

**Table 1.** CT image acquisition protocols.

were subsequently selected for downstream analysis, as they collectively captured the majority of variance while minimizing redundancy. ANOVA was used to select the PCA features for training a Categorical Boosting (CatBoost) model. In this study, due to the limited size of the test set which may not fully represent the model's performance, we performed 5-fold cross-validation on the training set to more robustly evaluate the model's generalization capability while maximizing the use of the available data. After the CatBoost model was trained in the training dataset, the cases in the test dataset were predicted by the model. The probability predicted by CatBoost was used to evaluate the efficacy of the radiomics model.

Evaluation of the model

Model evaluation was performed using the R programming language (version 4.1.1). In the test cohort, the predictive efficacy of the radiomics model and the CT impression were assessed using receiver operating characteristic (ROC) curves, generated with the pROC package. Precision-recall (PR) curves and calibration curves (bootstrapped with 1,000 repetitions for bias correction) were plotted using the PRROC and rms packages, respectively. Model performance metrics—including accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR), and negative likelihood ratio (NLR)—were calculated using the modEvA package.

Statistical analysis

The statistical analysis was conducted using the R software, version 4.1.3. For continuous variables, those following a normal distribution were presented as mean ± standard deviation, while non-normally distributed variables were depicted as median [1st quartile, 3rd quartile]. Categorical variables were summarized as counts and percentages. Normality of distribution was assessed using the Kolmogorov-Smirnov test. Associations among categorical variables were evaluated using chi-square or Fisher's exact tests as appropriate. For two-group comparisons involving non-normally distributed data, the Mann-Whitney U test was utilized. In cases where multiple groups were compared and the data did not meet parametric assumptions, the Kruskal-Wallis test was applied. The DeLong test was employed to compare the areas under the Receiver Operating Characteristic (ROC) curves across various models. Differences were considered statistically significant at a P value of less than 0.05.

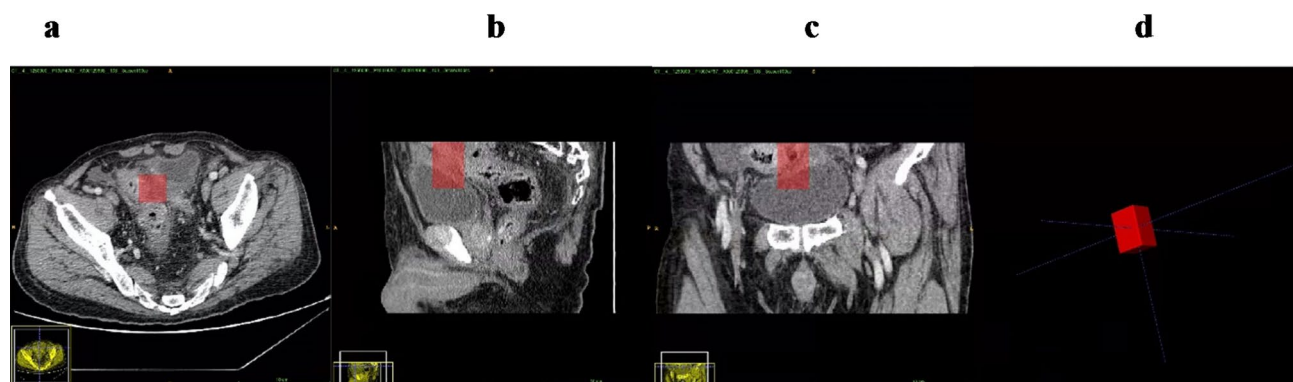
Results

Clinical characteristics of the cases

The clinical characteristics of the BI(+) and BI(-) patients are shown in Table 3. The characteristics of BI(+) and BI(-) patients were compared. No significant differences were found between the BI(+) and BI(-) groups for any of the clinical factors (all *P* > 0.05), except for the differentiation of the tumor in the test dataset (*P* = 0.018).

	Overall (n = 96)	Training dataset (n = 68)	Test dataset (n = 28)	P value
Age (yrs)				
Mean (SD)	60.4(12.3)	60.7(11.9)	59.8(13.3)	> 0.999
Gender				
Female	22 (22.9%)	16 (23.5%)	6 (21.4%)	> 0.999
Male	74 (77.1%)	52 (76.5%)	22 (78.6%)	
CEA				
< 5.0 ng/ml	75 (78.1%)	23 (82.1%)	52 (76.5%)	0.734
≥ 5.0 ng/ml	21 (21.9%)	5 (17.9%)	16 (23.5%)	
CA19-9				
< 37.0U/ml	56 (58.3%)	10 (35.7%)	46 (67.6%)	0.008
≥ 37.0U/ml	40 (41.7%)	18 (64.3%)	22 (32.4%)	
BMI				
Mean (SD)	23.2 (3.46)	22.2 (2.68)	23.6 (3.68)	0.050
Location of the tumor				
Descending colon	1 (1.0%)	0 (0%)	1 (3.6%)	0.354
Sigmoid colon	60 (62.5%)	42 (61.8%)	18 (64.3%)	
Rectum	33 (34.4%)	25 (36.8%)	8 (28.6%)	
Sigmoid colon and rectum	2 (2.1%)	1 (1.5%)	1 (3.6%)	
Differentiation				
Low	5 (5.2%)	3 (4.4%)	2 (7.1%)	0.250
Middle	78 (81.3%)	55 (80.9%)	23 (82.1%)	
High	3 (3.1%)	1 (1.5%)	2 (7.1%)	
Mucinous	10 (10.4%)	9 (13.2%)	1 (3.6%)	
Diameter (cm)				
Median [Q1,Q3]	7.00 [5.50,9.00]	7.50 [5.88,9.00]	6.25 [5.00,7.63]	0.051

**Table 2.** Clinical characteristics of the training and test datasets.



**Fig. 2.** Region of interest of the models. A rectangle-shaped ROI was annotated at the area where the bladder and the CRC were closest. The ROIs are shown in the axial (a), sagittal (b), and coronal (c) planes and were viewed three-dimensionally (d).

### Results of the CT impression

The results of the CT impression are shown in Table 4. The inter-reader agreement for the CT image features of CRC expansion, bladder wall thickening, and perivesical stranding was 0.876, 0.761, and 0.769, respectively. In the training dataset, the diagnostic accuracy of the CT impression was 64.7%. In the test dataset, the diagnostic accuracy of the CT impression was 53.6%.

### Results of the radiomics model

After PCC analysis, 88 features were selected as robust features and passed for PCA. Then, feature reduction was performed by PCA (Fig. 3; Table 5), and five PCA features were selected to train the CatBoost model. For training the CatBoost model, accuracy was used to select the optimal model using the largest value. The final values used for the model were depth = 2, learning\_rate = 0.1, iterations = 100, l2\_leaf\_reg = 1e-06, rsm = 0.9,

	Training dataset			Test dataset		
	BI(-)	BI(+)		BI(-)	BI(+)	
	(N = 42)	(N = 26)		(N = 19)	(N = 9)	
Gender						
Male	30 (71.4%)	22 (84.6%)	0.341	14 (73.7%)	8 (88.9%)	0.673
Female	12 (28.6%)	4 (15.4%)		5 (26.3%)	1 (11.1%)	
Age						
Mean (SD)	61.8 (12.5)	59.0 (10.9)	0.253	59.8 (12.2)	59.6 (16.2)	0.921
CEA						
< 5.0 ng/ml	33 (78.6%)	19 (73.1%)	0.822	15 (78.9%)	8 (88.9%)	0.910
≥ 5.0 ng/ml	9 (21.4%)	7 (26.9%)		4 (21.1%)	1 (11.1%)	
CA19-9						
< 37.0U/ml	28 (66.7%)	18 (69.2%)	> 0.999	8 (42.1%)	2 (22.2%)	0.546
≥ 37.0U/ml	14 (33.3%)	8 (30.8%)		11 (57.9%)	7 (77.8%)	
BMI						
Mean (SD)	23.9 (3.64)	22.9 (3.73)	0.218	22.8 (2.19)	20.9 (3.25)	0.085
Location						
Descending colon	0 (0%)	0 (0%)	0.683	1 (5.3%)	0 (0%)	0.426
Sigmoid colon	25 (59.5%)	17 (65.4%)		12 (63.2%)	6 (66.7%)	
Rectum	16 (38.1%)	9 (34.6%)		6 (31.6%)	2 (22.2%)	
Sigmoid colon and rectum	1 (2.4%)	0 (0%)		0 (0%)	1 (11.1%)	
Differentiation						
Low	1 (2.4%)	2 (7.7%)	0.629	0 (0%)	2 (22.2%)	0.018
Middle	34 (81.0%)	21 (80.8%)		18 (94.7%)	5 (55.6%)	
High	1 (2.4%)	0 (0%)		0 (0%)	2 (22.2%)	
Mucinous	6 (14.3%)	3 (11.5%)		1 (5.3%)	0 (0%)	
Diameter						
Median [Min, Max]	7.25 [3.00, 17.0]	7.75 [3.00, 14.0]	0.626	6.00 [4.00, 13.0]	7.50 [4.00, 13.0]	0.729

**Table 3.** Clinical characteristics of patients with and without BI.

	Training dataset		Test dataset	
	CT impression (+)	CT impression (-)	CT impression (+)	CT impression (-)
BI(+)	21	5	6	3
BI(-)	19	23	10	9

**Table 4.** Results of the CT impression.

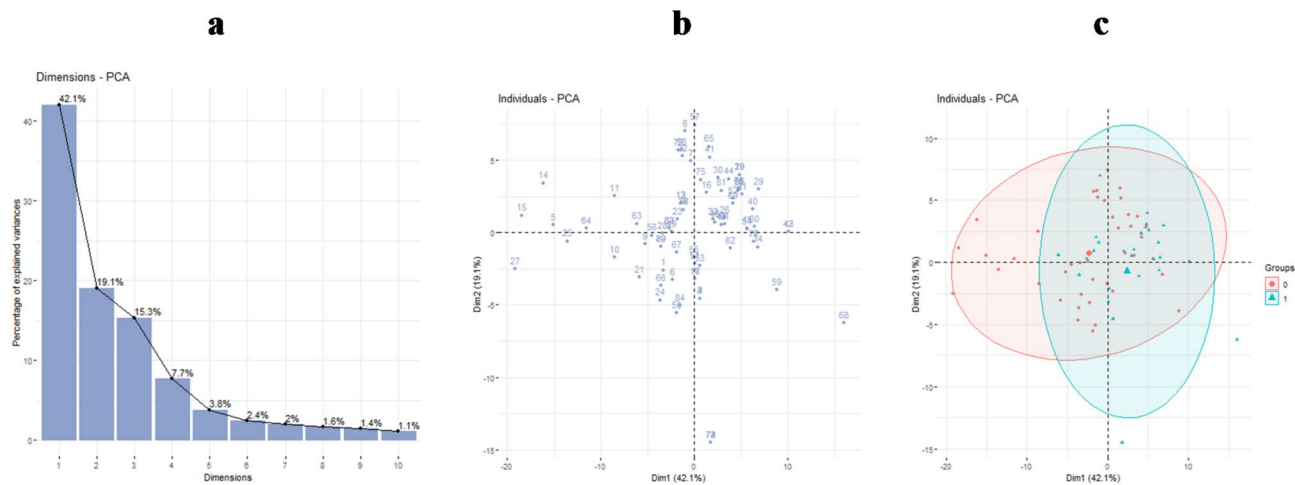
and border\_count = 255. The relative importance of the PCA features was 24.12%, 21.17%, 20.73%, 19.50%, and 14.49% in the overall model. The AUC values and their corresponding 95% confidence intervals from the 5-fold cross-validation were 0.706 (0.551–0.806), 0.797 (0.636–0.891), 0.747 (0.588–0.799), 0.831 (0.774–0.858), and 0.891 (0.768–0.991), respectively.

**Evaluation of the model**

The area under the curve (AUC) and other evaluation metrics for CT impression and the radiomics model are shown in Table 6; Fig. 4, including AUC, accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR), and negative likelihood ratio (NLR), respectively. In the training dataset, the AUC of the radiomic model [0.864 (95% CI: 0.778, 0.951)] was significantly greater than that of CT impression [0.678 (95% CI: 0.569, 0.786), *P* = 0.007]. In the test dataset, the AUC of the radiomic model [0.883 (95% CI: 0.699, 1.000)] was also significantly greater than that of CT impression [0.570 (95% CI: 0.370, 0.770), *P* = 0.040].

The PR curves of the CT impression and the radiomics model are shown in Fig. 5. The calibration curves are shown in Fig. 6. With 1,000 repetitions of bootstrapping, the calibration curve of the CT impression yielded a mean absolute error of 0.042, a mean squared error of 0.002, and a 0.9 quantile of absolute error of 0.043. The calibration curve of the radiomics model yielded a mean absolute error of 0.029, a mean squared error of 0.001, and a 0.9 quantile of absolute error of 0.071.





**Fig. 3.** Feature reduction was performed by PCA. **(a)** Variance against the number of dimensions in the result of PCA. **(b)** Individual case in the PCA. Each individual case is plotted on the platform of dimension 1 (Dim1) and dimension 2 (Dim2). Dim1 and Dim2 contribute to the principal component with proportions of 42.1% and 19.1%, respectively. **(c)** Individual case in the PCA with different groups. Group 0 [BI(-)] in red and group 1 [BI(+)] in green are plotted in the platform of dimension 1 (Dim1) and dimension 2 (Dim2).

	PC1	PC2	PC3	PC4	PC5
Standard deviation	6.0843	4.0977	3.6719	2.59871	1.83007
Proportion of variance	0.4207	0.1908	0.1532	0.07674	0.03806
Cumulative proportion	0.4207	0.6115	0.7647	0.84143	0.87949

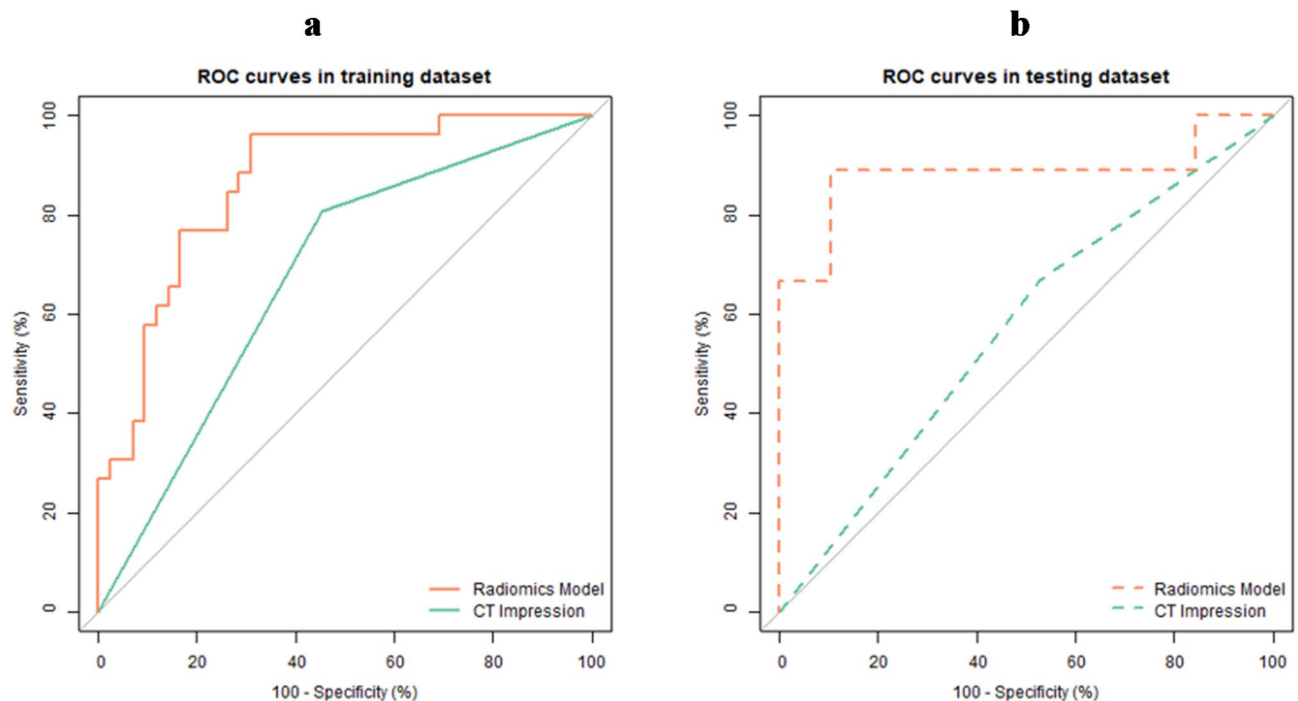
**Table 5.** Standard deviation and proportion of the top 5 PCA features.

	CT impression	Radiomic model
AUC	0.678 (0.569, 0.786)	0.864 (0.778, 0.951)
ACC	0.647 (0.640, 0.654)	0.794 (0.789, 0.799)
SEN	0.808 (0.656, 0.959)	0.962 (0.888, 1.000)
SPE	0.548 (0.397, 0.698)	0.690 (0.551, 0.830)
PPV	0.525 (0.370, 0.680)	0.658 (0.507, 0.809)
NPV	0.821 (0.680, 0.963)	0.967 (0.902, 1.031)
PLR	1.785 (1.219, 2.616)	3.107 (1.965, 4.912)
NLR	0.351 (0.152, 0.809)	0.056 (0.008, 0.385)

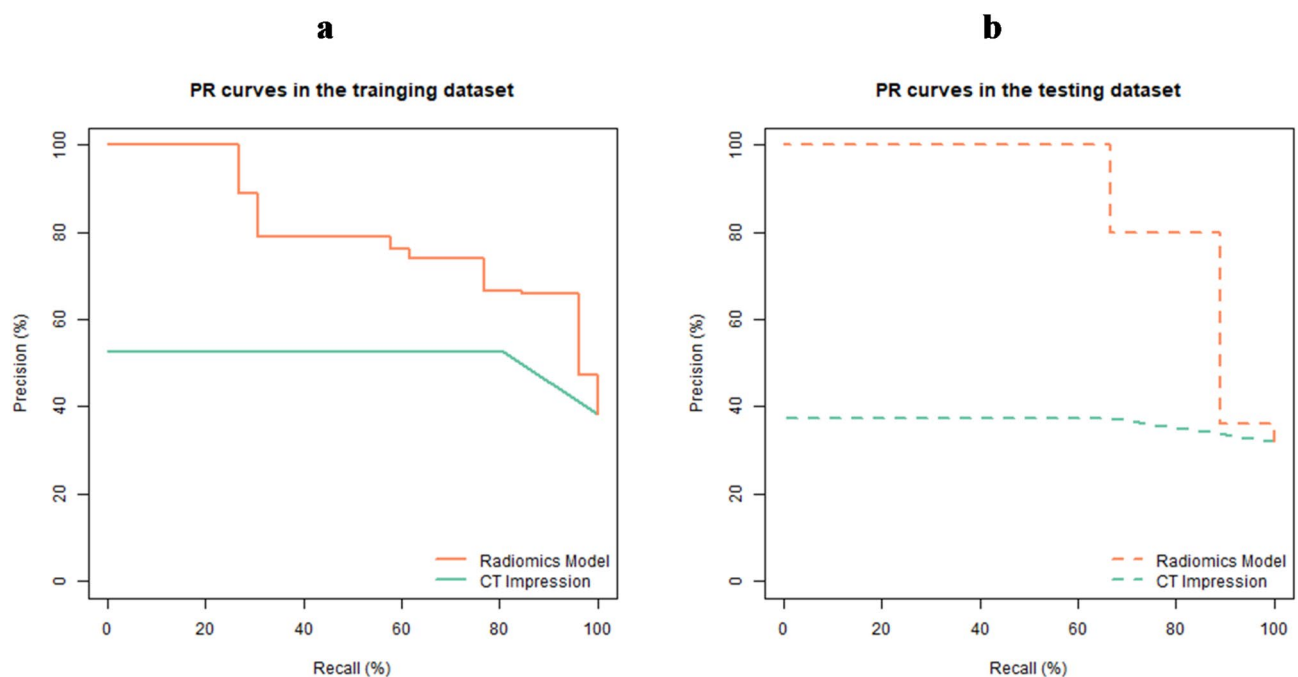
**Table 6.** AUCs and other evaluation metrics for the CT impression and the radiomics model.

Discussion

Predicting BI by locally advanced CRC is important because it can help determine the stage of the cancer and guide treatment decisions. Studies have shown that the most relevant preoperative predictors of definite BI in patients with CRC are gross hematuria, a tumor visible during cystoscopy, and abnormal CT findings<sup>5</sup>. CT can provide detailed images of the bladder and surrounding structures, allowing radiologists to identify abnormal findings such as gross tumor invasion, an enhancing mass at the bladder wall, irregular bladder mucosa, and loss of the perivesical fat plane that may indicate cancer. However, CT scans are not definitive and are often used in combination with other diagnostic methods, such as cystoscopy and MRI, to confirm the diagnosis. Relatively low concordance rates, approximately 35–55%, between preoperative CT scans and final histopathology results have been reported<sup>8,18</sup>, which means that there may be excessive preoperative staging, which leads to an increase in the proportion of neoadjuvant therapy and the expansion of the scope of intraoperative resection. Thus, improving the coincidence rate of imaging and pathology is the future direction of diagnosis. Radiomics is a field of medical imaging in which advanced computational methods are used to extract a large number of quantitative features from medical images. In the case of CRC, radiomics has been shown to be useful for evaluating the aggressiveness of the disease, as well as for predicting patient outcomes<sup>19</sup>.



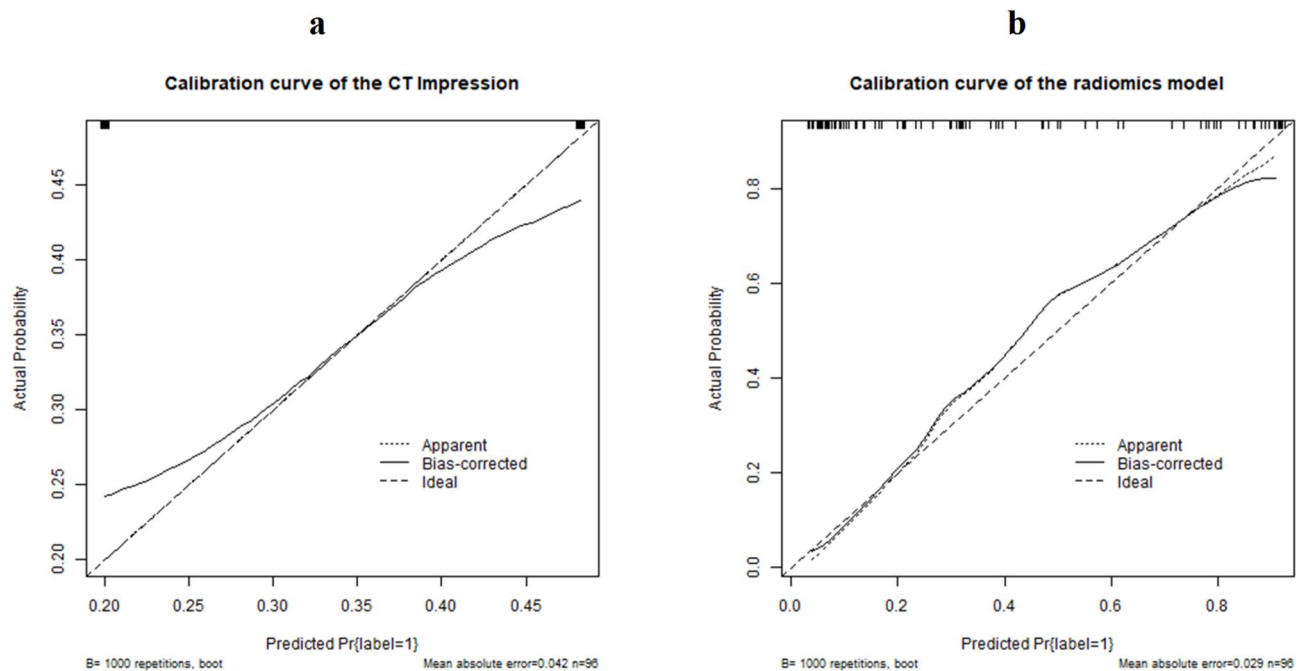
**Fig. 4.** ROC curves of the CT impression and the radiomic model in the training dataset (a), the AUC of the radiomic model [0.864 (95% CI: 0.778, 0.951)] was significantly greater than that of the CT impression [0.678 (95% CI: 0.569, 0.786),  $P = 0.007$ ]. In the test dataset (b), The AUC of the radiomic model [0.883 (95% CI: 0.699, 1.000)] was significantly greater than that of the CT impression [0.570 (95% CI: 0.370, 0.770),  $P = 0.040$ ].



**Fig. 5.** Precision-recall curves of the CT impression and the radiomic model in the training dataset (a) and the test dataset (b).

In comparison to prior radiomics studies on colorectal cancer, our work advances the field in two key aspects. First, while existing CT-based radiomics research in CRC has addressed diverse diagnostic tasks (e.g., tumor staging, lymph node metastasis), our study is the first to specifically evaluate bladder invasion. This addresses a critical unmet need in preoperative assessment, where distinguishing true BI from inflammatory





**Fig. 6.** Calibration curve of the CT impression (a), with 1,000 repetitions of bootstrapping, the calibration curve yielded a mean absolute error of 0.042, a mean squared error of 0.002, and a 0.9 quantile of absolute error of 0.043. Calibration curve of the radiomics model (b), with 1,000 repetitions of bootstrapping, the calibration curve yielded a mean absolute error of 0.029, a mean squared error of 0.001, and a 0.9 quantile of absolute error of 0.071.

adhesions remains a major clinical challenge. By resolving this ambiguity, our model could reduce unnecessary multivisceral resections and improve surgical planning. Second, we introduce a standardized rectangular ROI spanning the tumor-bladder interface. Unlike conventional approaches that rely on irregular tumor contours, this method ensures consistent, reproducible annotations across radiologists, thereby enhancing the validity and reliability of radiomics feature extraction. Together, these advances strengthen the clinical utility of radiomics for guiding preoperative surgical decisions in CRC.

In this study, we reviewed 96 CRC patients who underwent surgery for suspected bladder invasion. A CT visual assessment model and a radiomics model were developed and compared. We found that there were no significant differences in age, sex, BMI, CA19-9, CEA, tumor location, or tumor diameter between patients with and without BI. However, there was a statistically significant difference in tumor differentiation between the two groups in the test dataset. In the test dataset, the AUC of the radiomics model (0.883) was significantly greater than that of the CT visual assessment (0.570,  $P = 0.040$ ).

Accurate and consistent annotation of ROIs is crucial for our radiomics study to be robust<sup>20</sup>. Our PubMed search results indicate that most radiomics studies on rectal cancer use manual segmentation, in which radiologists annotate the location and precise boundary of the tumor<sup>21</sup>. In this study, we chose the ROI as the area where the bladder and the CRC were closest because we believe that this area contains important information about the biology of the tumor, such as the aggressiveness of the tumor, the extent of invasion, and the potential for metastasis. Using this clear definition of the ROI, we can ensure that the radiologists consistently annotate the ROI, making the results of the study valid and reliable. To evaluate the bias in the ROI, we enlisted the participation of two radiologists with varying levels of experience and assessed the intra- and interrater repeatability of the annotations. The results demonstrated good consistency between them.

At present, comprehensive treatment based on surgical resection is still the treatment mode for locally advanced CRC with bladder involvement. This study revealed that applying radiomics models for the prediction of BI by CRC is feasible and may perform better than human radiologists, expecting to improve the accuracy of clinical staging of locally advanced CRC, guide the selection of appropriate treatment strategies and decrease unnecessary bladder resection.

A first diagnosis of locally advanced CRC with suspected bladder infiltration is recommended for radiomics assessment. For locally advanced colon cancer, if the bladder is assessed as non-invasive using the radiomics model, surgical exploration may be an option for attempting to preserve the bladder intactly or reduce the extent of cystectomy; otherwise, combined partial or total cystectomy or neoadjuvant therapy may be an option. For locally advanced rectal cancer, if the bladder is assessed as non-invasive by the radiomics model, attempts may be made to preserve the bladder intactly or reduce the extent of cystectomy after neoadjuvant therapy.

This study has several limitations. First, the small sample size (particularly the limited number of positive invasion cases,  $n = 8$  in the test set) may compromise the statistical power and robustness of the radiomics model, thereby affecting its reproducibility—a critical criterion for radiomics models. Second, the use of heterogeneous

CT scanners with varying imaging protocols could introduce variability in radiomic feature extraction, further limiting the reproducibility of results. Future work should incorporate harmonization techniques, such as batch effect correction using empirical Bayes methods (e.g., ComBat) or deep learning-based domain adaptation, to mitigate scanner- and parameter-related variability. Third, the study population was primarily composed of patients who underwent CRC resection and cystectomy, which may introduce selection bias as conservatively managed patients were excluded. Future prospective cohorts including both surgical and non-surgical patients are needed to better reflect real-world clinical diversity. Fourth, the single-center design and limited dataset restrict the generalizability of our findings; external validation with multicenter data is essential to confirm the model's applicability across diverse clinical settings. Finally, manual ROI annotation, while necessary for this study, is labor-intensive and prone to inter-observer variability. Automated segmentation methods should be developed in future work to improve efficiency and consistency.

## Conclusions

In summary, based on the results of this study, it appears that using radiomics models for the prediction of BI by locally advanced CRC is feasible and may perform better than human radiologists. The application of radiomic models can assist in the preoperative staging of locally advanced CRC, guide the selection of treatment strategies, and reduce unnecessary bladder resection.

## Data availability

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

Received: 6 October 2024; Accepted: 17 April 2025

Published online: 02 May 2025

## References

- Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. <https://doi.org/10.3322/caac.21492> (2018).
- Gebhardt, C., Meyer, W., Ruckriegel, S. & Meier, U. Multivisceral resection of advanced colorectal carcinoma. *Langenbecks Arch. Surg.* **384**, 194–199. <https://doi.org/10.1007/s004230050191> (1999).
- Chen, Y. G., Liu, Y. L., Jiang, S. X. & Wang, X. S. Adhesion pattern and prognosis studies of T4N0M0 colorectal cancer following En bloc multivisceral resection: Evaluation of T4 subclassification. *Cell Biochem. Biophys.* **59**, 1–6. <https://doi.org/10.1007/s12013-010-9106-z> (2011).
- Kobayashi, T. et al. High incidence of urinary bladder involvement in carcinomas of the sigmoid and rectum: a retrospective review of 580 patients with colorectal carcinoma. *J. Surg. Oncol.* **84**, 209–214. <https://doi.org/10.1002/jso.10322> (2003).
- Woranisarakul, V. et al. Accuracy of preoperative urinary symptoms, urinalysis, computed tomography and cystoscopic findings for the diagnosis of urinary bladder invasion in patients with colorectal cancer. *Asian Pac. J. Cancer Prev.* **15**, 7241–7244. <https://doi.org/10.7314/apjcp.2014.15.17.7241> (2014).
- Lehnert, T. et al. Multivisceral resection for locally advanced primary colon and rectal cancer: An analysis of prognostic factors in 201 patients. *Ann. Surg.* **235**, 217–225. <https://doi.org/10.1097/0000658-200202000-00009> (2002).
- Nakafusa, Y. et al. Comparison of multivisceral resection and standard operation for locally advanced colorectal cancer: analysis of prognostic factors for short-term and long-term outcome. *Dis. Colon Rectum.* **47**, 2055–2063. <https://doi.org/10.1007/s10350-004-0716-7> (2004).
- Mohan, H. M., Evans, M. D., Larkin, J. O., Beynon, J. & Winter, D. C. Multivisceral resection in colorectal cancer: A systematic review. *Ann. Surg. Oncol.* **20**, 2929–2936. <https://doi.org/10.1245/s10434-013-2967-9> (2013).
- Brown, K. G. et al. Urological complications after cystectomy as part of pelvic exenteration are higher than that after cystectomy for primary bladder malignancy. *J. Surg. Oncol.* **115**, 307–311. <https://doi.org/10.1002/jso.24511> (2017).
- Nakafusa, Y. et al. Modification of lymph node dissection for colon cancer by clinical diagnosis. *Hepatogastroenterology* **51**, 722–726 (2004).
- Hunter, J. A., Ryan, J. A. Jr. & Schultz, P. En bloc resection of colon cancer adherent to other organs. *Am. J. Surg.* **154**, 67–71. [https://doi.org/10.1016/0002-9610\(87\)90292-3](https://doi.org/10.1016/0002-9610(87)90292-3) (1987).
- Montesani, C. et al. Extended resection in the treatment of colorectal cancer. *Int. J. Colorectal Dis.* **6**, 161–164. <https://doi.org/10.1007/bf00341238> (1991).
- Lopez, M. J. Multivisceral resections for colorectal cancer. *J. Surg. Oncol.* **76**, 1–5 (2001).
- Visokai, V. et al. Multiorgan resections for advanced colorectal cancer. *Anticancer Res.* **26**, 3183–3186 (2006).
- Izbicki, J. R. et al. Extended resections are beneficial for patients with locally advanced colorectal cancer. *Dis. Colon Rectum.* **38**, 1251–1256. <https://doi.org/10.1007/bf02049148> (1995).
- Aerts, H. J. et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **5**, 4006. <https://doi.org/10.1038/ncomms5006> (2014).
- Yip, S. S. & Aerts, H. J. Applications and limitations of radiomics. *Phys. Med. Biol.* **61**, R150–166. <https://doi.org/10.1088/0031-9155/61/13/r150> (2016).
- Yoshida, T., Shida, D., Taniguchi, H., Tsukamoto, S. & Kanemitsu, Y. Long-Term outcomes following partial versus complete cystectomy in advanced colorectal Cancer with regarding to the extent of bladder invasion. *Ann. Surg. Oncol.* **26**, 1569–1576. <https://doi.org/10.1245/s10434-019-07276-0> (2019).
- Inchingolo, R. et al. Radiomics in colorectal cancer patients. *World J. Gastroenterol.* **29**, 2888–2904. <https://doi.org/10.3748/wjg.v29.i19.2888> (2023).
- Horvat, N., Bates, D. D. B. & Petkovska, I. Novel imaging techniques of rectal cancer: What do radiomics and radiogenomics have to offer? A literature review. *Abdom. Radiol.* **44**, 3764–3774. <https://doi.org/10.1007/s00261-019-02042-y> (2019).
- Hou, M. & Sun, J. H. Emerging applications of radiomics in rectal cancer: State of the art and future perspectives. *World J. Gastroenterol.* **27**, 3802–3814. <https://doi.org/10.3748/wjg.v27.i25.3802> (2021).

## Acknowledgements

The authors acknowledge Yaofeng Zhang and Xiangpeng Wang from Beijing Smart Tree Medical Technology for technical support.

### Author contributions

X.Y.W. and X.W. were the major contributors to the design of this study and revised the manuscript. J.G.W. and K.X.W. collected the clinical information, processed the imaging data, analyzed the data, and wrote the manuscript. J.L.Z., Y.C.W., Y.J., G.W.C., Z.B.L., T.W., and Y.L.W. participated in the design of the study and provided clinical guidance. All the authors read and approved the final manuscript.

### Declarations

#### Competing interests

The authors declare no competing interests.

#### Ethical approval and consent to participate

This retrospective study was approved by the local Institutional Review Board (Peking University First Hospital 2019–170), and the requirement for written consent was waived.

### Additional information

**Correspondence** and requests for materials should be addressed to X.W. or X.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025