Biomedical Physics & Engineering Express

PAPER

Multimodality radiomics prediction of radiotherapy-induced the early proctitis and cystitis in rectal cancer patients: a machine learning study

RECEIVED
24 July 2023

REVISED
8 November 2023

ACCEPTED FOR PUBLICATION
23 November 2023

PUBLISHED
20 December 2023

Samira Abbaspour[1,2], **Maedeh Barahman**[3], **Hamid Abdollahi**[4,5], **Hossein Arabalibeik**[6], **Ghasem Hajainfar**[7], **Mohammadreza Babaei**[8], **Hamed Iraji**[8], **Mohammadreza Barzegartahamtan**[9], **Mohammad Reza Ay**[1,2,*] and **Seied Rabi Mahdavi**[10,11,*]

[1] Department of Medical Physics and Biomedical Engineering, Tehran University of Medical Sciences, Tehran, Iran
[2] Research Center for Molecular and Cellular Imaging, Advanced Medical Technologies and Equipment Institute (AMTEI), Imam Khomeini Hospital, Tehran University of Medical Sciences, Tehran, Iran
[3] Department of Radiation Oncology, Firoozgar Hospital, Firoozgar Clinical Research Development Center (FCRDC), Iran University of Medical Sciences, Tehran, Iran
[4] Department of Radiologic Sciences and Medical Physics, Kerman University of Medical Sciences, Kerman, Iran
[5] Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC, Canada
[6] Research Center for Science and Technology in Medicine (RCSTM), Tehran University of Medical Sciences, Tehran, Iran
[7] Rajaie Cardiovascular Medical & Research Center, Iran University of Medical Science, Tehran, Iran
[8] Department of Interventional Radiology, Firouzgar Hospital, Iran University of Medical Sciences, Tehran, Iran
[9] Clinical Research Development Unit, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[10] Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[11] Radiation Biology Research Center, Iran University of Medical Sciences, Tehran, Iran
[*] Authors to whom any correspondence should be addressed.

E-mail: mohammadreza_ay@tums.ac.ir **and** srmahdavi@hotmail.com

Keywords: proctitis, cystitis, rectal cancer, radiomics features, machine learning

Supplementary material for this article is available online

Abstract

Purpose. This study aims to predict radiotherapy-induced rectal and bladder toxicity using computed tomography (CT) and magnetic resonance imaging (MRI) radiomics features in combination with clinical and dosimetric features in rectal cancer patients. *Methods.* A total of sixty-three patients with locally advanced rectal cancer who underwent three-dimensional conformal radiation therapy (3D-CRT) were included in this study. Radiomics features were extracted from the rectum and bladder walls in pretreatment CT and MR-T2W-weighted images. Feature selection was performed using various methods, including Least Absolute Shrinkage and Selection Operator (Lasso), Minimum Redundancy Maximum Relevance (MRMR), Chi-square (Chi2), Analysis of Variance (ANOVA), Recursive Feature Elimination (RFE), and SelectPercentile. Predictive modeling was carried out using machine learning algorithms, such as K-nearest neighbor (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Naive Bayes (NB), Gradient Boosting (XGB), and Linear Discriminant Analysis (LDA). The impact of the Laplacian of Gaussian (LoG) filter was investigated with sigma values ranging from 0.5 to 2. Model performance was evaluated in terms of the area under the receiver operating characteristic curve (AUC), accuracy, precision, sensitivity, and specificity. *Results.* A total of 479 radiomics features were extracted, and 59 features were selected. The pre-MRI T2W model exhibited the highest predictive performance with an AUC: 91.0/96.57%, accuracy: 90.38/96.92%, precision: 90.0/97.14%, sensitivity: 93.33/96.50%, and specificity: 88.09/97.14%. These results were achieved with both original image and LoG filter (sigma = 0.5–1.5) based on LDA/DT-RF classifiers for proctitis and cystitis, respectively. Furthermore, for the CT data, AUC: 90.71/96.0%, accuracy: 90.0/96.92%, precision: 88.14/97.14%, sensitivity: 93.0/96.0%, and specificity: 88.09/97.14% were acquired. The highest values were achieved using XGB/DT-XGB classifiers for proctitis and cystitis with LoG filter (sigma = 2)/LoG

filter (sigma = 0.5–2), respectively. MRMR/RFE-Chi2 feature selection methods demonstrated the best performance for proctitis and cystitis in the pre-MRI T2W model. MRMR/MRMR-Lasso yielded the highest model performance for CT. *Conclusion.* Radiomics features extracted from pretreatment CT and MR images can effectively predict radiation-induced proctitis and cystitis. The study found that LDA, DT, RF, and XGB classifiers, combined with MRMR, RFE, Chi2, and Lasso feature selection algorithms, along with the LoG filter, offer strong predictive performance. With the inclusion of a larger training dataset, these models can be valuable tools for personalized radiotherapy decision-making.

1. Introduction

Rectal cancer is the second most prevalent form of cancer in the large intestine, and its primary treatment modalities include radiotherapy, chemotherapy, and surgery [1, 2]. Radiotherapy, in particular, can lead to various early and late effects on the rectum and bladder, potentially causing proctitis, cystitis, and a range of toxicities such as bowel obstruction, fistula, perforation, dysuria, hematuria, and a significant decrease in quality of life. These complications are contingent on multiple clinical factors, including radiation dose and the patient's clinical, biological, and genomic characteristics [3–5].

1.1. The potential of radiomics

The analysis of radiation toxicity in rectal cancer can be enhanced through the use of image-based features, which aid physicians in mitigating radiation risks and determining the feasibility of local tumor control [6, 7]. Radiomics, a novel imaging analysis approach, involves the quantification of high-dimensional data extracted from medical images, providing valuable information about pathophysiological properties [8–10]. In the context of radiotherapy, radiomics feature analysis of the target volume and organs at risk (OARs) can have various applications, such as diagnostics, risk stratification, disease-free survival prediction, automatic segmentation, target volume definition, toxicity prognosis, treatment plan optimization, adaptive re-planning, decision support, treatment response assessment, and follow-up [8, 11–13].

1.2. Predictive modeling for radiation-induced damage

Developing predictive models for radiation-induced toxicity is crucial for optimizing radiotherapy dosages and enhancing patients' quality of life [14]. Artificial intelligence can predict potential complications in the rectum and bladder, such as proctitis and cystitis, by learning from previous cases [15]. If reliable knowledge can be effectively extracted, patient characteristics, imaging data, and planned radiation therapy can be correlated with the likelihood of severe symptoms. This correlation enables the creation of a classification model capable of identifying at-risk patients [16]. Indeed, multiple studies have highlighted the utility of radiomics analysis in quantifying radiation therapy-induced damage in various organs, including the bladder, rectum, parotid, and lung [6, 17–22].

1.3. The current study

While a few studies have reported image-based predictors for proctitis and cystitis [4, 14, 17, 18, 23], the combination of different feature types has been explored infrequently [24, 25]. In this study, we propose a scheme for predicting rectal and bladder toxicity using a combination of clinical, dosimetric, and radiomics features extracted from pretreatment planning computed tomography (CT) and magnetic resonance imaging (MRI) in rectal cancer patients. Our goal is to classify patients at risk of developing proctitis and cystitis using machine learning methods.

2. Methods

Figure 1 provides an overview of the methods employed in this study, which can be divided into two main phases: Data Acquisition, and Model Building and Evaluation. Each phase is explained in detail separately.

2.1. Data collection

The study received approval from the National Ethical Committee (registration No. IR.TUMS.MEDICINE.REC.1399.244). Informed consent was secured from all patients for the use of their clinical data in the current project. The study included sixty-three patients with locally advanced rectal tumors who met the following criteria: primary rectal cancer confirmed through biopsy, TNM stage (cT3–4) or lymph node involvement, and neoadjuvant chemoradiotherapy followed by surgery (the standard method) [26]. Exclusion criteria encompassed a history of prior chemotherapy or radiotherapy in the pelvic region, distant metastases (which require different, palliative treatment), and an absence of follow-up data.

2.2. Clinical and dosimetric data

Several clinical and dosimetric parameters were examined as potential predictors of rectal and bladder toxicity. Patient demographic and clinical information, including sex, age, baseline symptoms before radiotherapy, smoking history, TNM staging, carcinoembryonic antigen (CEA), tumor length (the

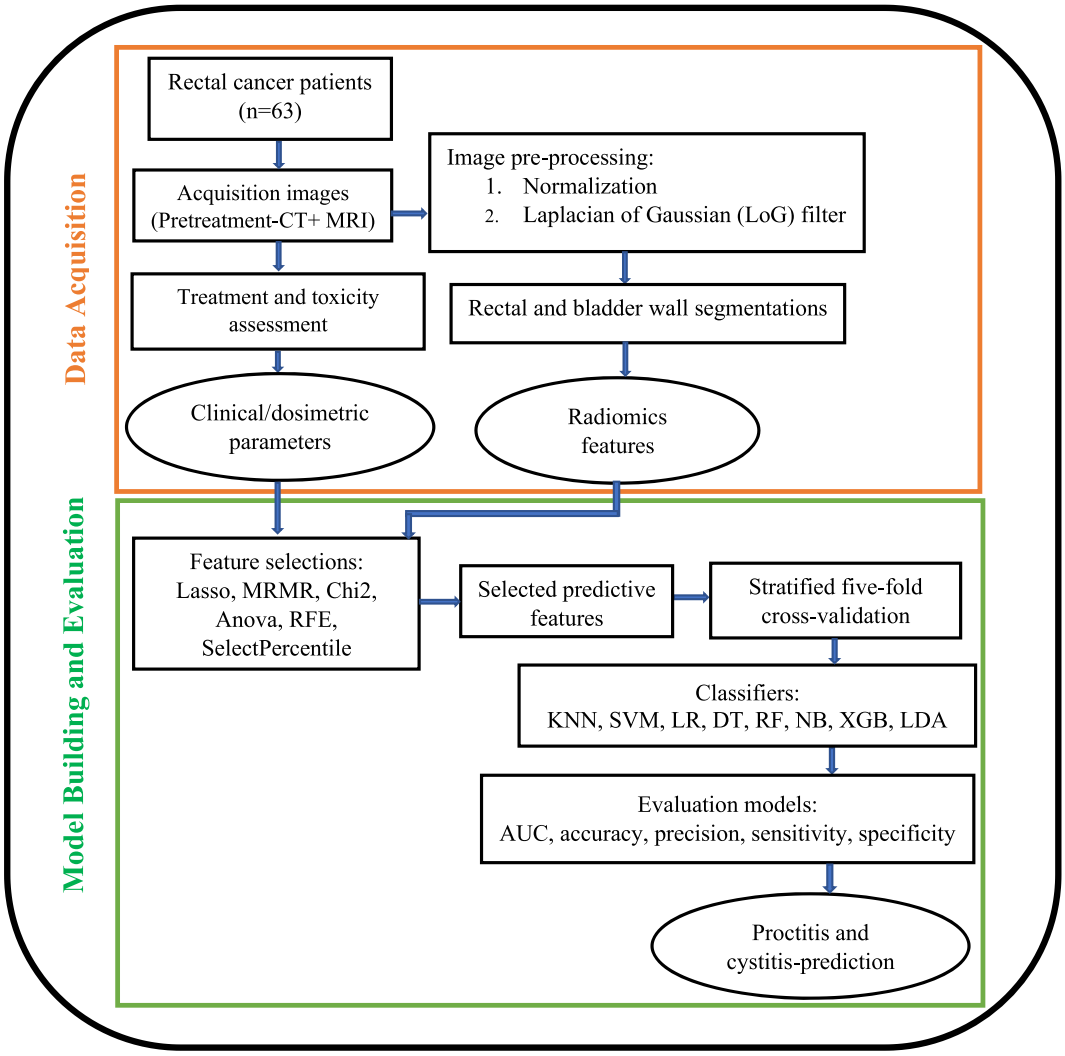

Figure 1. Flowchart of the study.

distance of the tumor from the anal verge as starting point to the end of the tumor, measured by an expert physician using available tools), and tumor differentiation (well, moderate, and poor), was gathered from electronic medical records. In addition, dosimetric data derived from dose-volume histograms (DVH) were recorded, including the average dose (D mean), minimum dose (D min), maximum dose (D max), and the dose reaching 10%–100% of the desired volume (D10 to D100) for the rectum and bladder walls.

2.3. CT and MRI protocols

Planning CT images were acquired from the patients prior to radiotherapy using a Siemens SOMATOM Sensation 64 machine with the following parameters: image matrices of 512 × 512 pixels, voxel size of $0.98 \times 0.98 \times 2.5\ mm^3$, kV of 120, mAs of 225, and a slice thickness of 1.5 mm. MR T2W images were obtained for all patients using a 1.5-Tesla Siemens Avanto machine, with a repetition time (TR) of 3000 ms, echo time (TE) of 101 ms, matrix size of 256 × 256, and a slice thickness of 3 mm.

2.4. Treatment and toxicity assessment

Patients received treatment through three-dimensional conformal radiation therapy (3D-CRT) utilizing an 18 MV linear accelerator (Siemens ONCOR Impression Plus, Germany). A total dose of 50.40 Gy, delivered in 28 fractions at 1.8 Gy per session, was administered for rectal cancer treatment. Chemotherapy consisted of oral capecitabine at a dose of 825 mg m^{-2}, taken twice daily during external radiation therapy. To assess toxicity in rectal cancer patients, the European Organization for the Research and Treatment of Cancer (EORTC) QLQ-C29 tool was employed. This tool includes 18 items that evaluate gastrointestinal symptoms, pain, and micturition problems. It also includes separate scales for participants with or without a stoma and specific items addressing sexual function for both men and women [27]. Proctitis and cystitis were considered as the study's endpoints. Patients were initially assessed for these toxicities before the commencement of radiotherapy, and further assessments were conducted during treatment and up to two months after

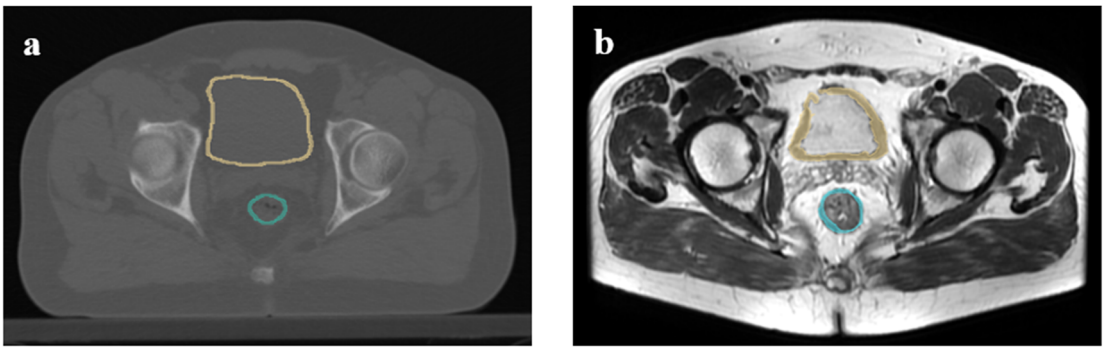

Figure 2. Segmentation of the rectum and bladder walls on (a) planning CT, (b) pre-MRI T2W of the rectal cancer patient in 3D-slicer.

completing chemotherapy (prior to surgery). Patients were categorized into two groups: those with complications scoring equal to or higher than 50 were labeled as 1, while those with scores lower than 50 were labeled as 0.

2.5. Image pre-processing

To minimize image noise and enhance sensitivity, as well as for discretization and the reduction of texture analysis variations, all image (CT/MRI) intensities were normalized within $\mu \pm 3\sigma$. Here, μ represents the mean value of gray levels within the region of interest (ROI), and σ represents the standard deviation.

A Laplacian of Gaussian (LoG) filter was employed to enhance performance, which combines a Gaussian smoothing operator, a standard deviation kernel, and an isotropic Laplacian filter. This filter is instrumental in highlighting image details at multiple scales [28–30]. In this study, LoG filters with sigma values of 0.5, 1, 1.5, and 2 were applied to the images.

2.6. Tumor segmentation

For the delineation of regions of interest (ROIs) on CT and MR T2W images, axial views were selected (figure 2). This segmentation process was conducted using the open-source 3D slicer software package (v. 4.10.2). The rectum and bladder walls were contoured by two experts, one radiologist and one radio-oncologist, who manually defined the ROIs for each image slice. Notably, the rectal lumen was excluded from the analysis, and only the portion of the rectum wall not encompassed by the gross tumor volume (GTV) was considered for contouring. The bladder and rectal wall voxel counts were found to be 5588 $\pm$ 87.55 and 2550 $\pm$ 75.30, respectively, with average volumes of 9434.57 $\pm$ 87.55 mm^3 and 5242.57 $\pm$ 75.30 mm^3. These values varied based on tumor size and patient anatomy.

2.7. Radiomics features

This study leveraged the pretreatment planning CT and MRI radiomics features extracted from the rectum and bladder walls to examine their correlation with radiation-induced toxicities. These radiomics features were obtained using the SlicerRadiomics extension of the 3D Slicer software, which incorporates the PyRadiomics library [31]. The radiomics features for the rectum and bladder walls were categorized into three groups: first-order, shape-based, and texture features. The texture features encompassed the following matrices: gray level run length matrix (GLRLM), gray level co-occurrence matrix (GLCM), gray level size zone matrix (GLSZM), gray level dependence matrix (GLDM), and neighboring gray tone difference matrix (NGTDM). These features were extracted from both original and filtered images, and additional details can be found in the Supplementary file (File 1).

Table 1. Feature selection and classification methods.

Feature selection method	Classification method
Lasso: Least Absolute Shrinkage and Selection Operator	**KNN:** K Nearest Neighbor
MRMR: Minimum Redundancy Maximum Relevance	**SVM:** Support Vector Machine
Chi2: Chi-Square	**LR:** Logistic Regression
Anova: Analysis of Variance	**DT:** Decision Tree
RFE: Recursive Feature Elimination	**RF:** Random Forest
SelectPersentile	**NB:** Naive Bayes
—	**XGB:** Extreme Gradient Boosting
—	**LDA:** Linear Discriminant Analysis

2.8. Feature selection and classification

To identify the most effective radiomics features for predicting proctitis and cystitis in rectal cancer patients, various feature selection methods were employed, as outlined in table 1. These selected radiomics features from the planning CT and MR images

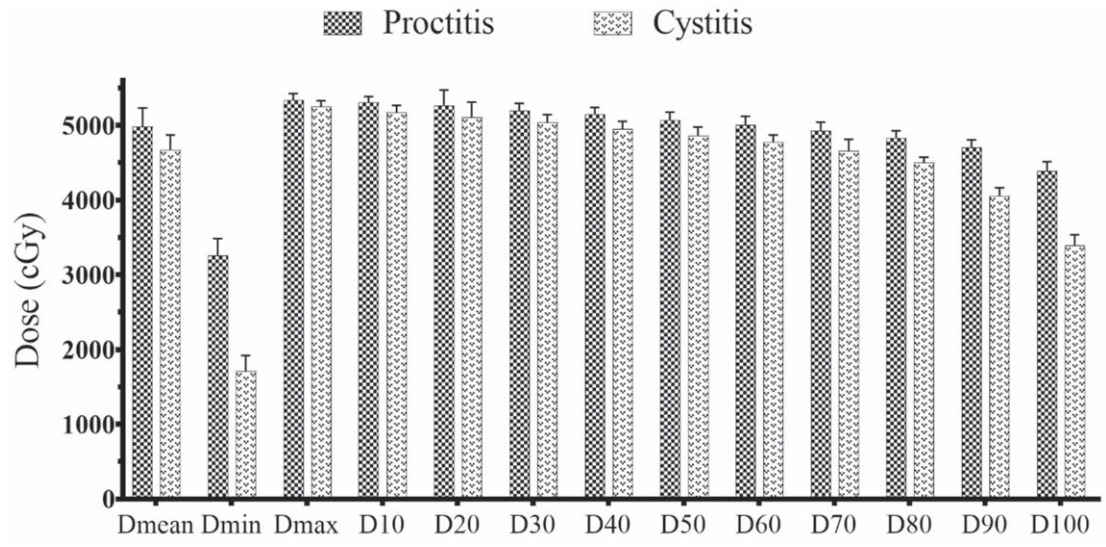

Figure 3. Comparison of the mean dosimetric parameters in patients with proctitis and cystitis toxicities.

Table 2. The patient demographic and clinical information.

Characteristic		All patients (n = 63) Value (%)	Proctitis (n = 48) Value (%)	Cystitis (n = 39) Value (%)
Age (years) Mean ± SD (range)	Male	62 ± 12.5 (29 – 81)	62 ± 11 (29 – 80)	60 ± 13 (31 – 81)
	Female	59 ± 12.5 (31 – 80)	61 ± 12 (32 – 80)	57 ± 11 (35 – 77)
Gender	Male	48 (76)	40 (83)	32 (82)
	Female	15 (24)	8 (17)	7 (18)
CEA	<5	29 (46)	37 (77)	30 (17)
	⩾5	34 (54)	11 (23)	9 (23)
T stage	T3	44 (70)	43 (90)	32 (82)
	T4	19(30)	5 (10)	7 (18)
N stage	N0	10 (16)	8 (17)	4 (10)
	N1	36 (57)	25 (52)	21(54)
	N2	17 (27)	15 (31)	14 (36)

were specifically related to the occurrence of proctitis and cystitis.

Subsequently, eight different classification methods, detailed in table 1, were utilized to construct prediction models. The stratified 5-fold cross-validation method [32, 33] was adopted, where all features were incorporated into the feature selection algorithms. A set of features was chosen for each fold. The selected features from the first fold were then input into the respective classifiers for that fold, and the evaluation parameters of the model were computed. This process was repeated five times, and the average values of the evaluation parameters were obtained for both the feature selection methods and classifiers.

For feature selection and classification, Python (v. 3.8) and the scikit-learn package (version 1.0.2) were employed.

2.9. Robust features against different delineations

Manual segmentation represents a common practice in the majority of radiomics studies, as it helps ensure consistency in ROI shapes, a critical factor affecting the reproducibility of predictive models. To assess the agreement between readers, the intra-class correlation coefficient (ICC), defined by equation (1), was used. ICC quantifies intra-reader agreement based on a two-way random effect model and has been widely employed to evaluate the reproducibility of radiomics features [34].

The calculations of ICC were conducted using R (v. 1.4.1106) and the IRR package (version 0.84.1). For the purpose of this study, a radiomics feature was considered reproducible if its ICC value was equal to or greater than 0.8 [30].

$$ICC = \frac{MS_R - MS_W}{MS_R + (k-1)MS_W} \quad (1)$$

where MS_R stands for 'the mean square for rows, each feature value for the two observers', MS_W refers to 'the mean square for the residual source of variance', and k is the number of observers.

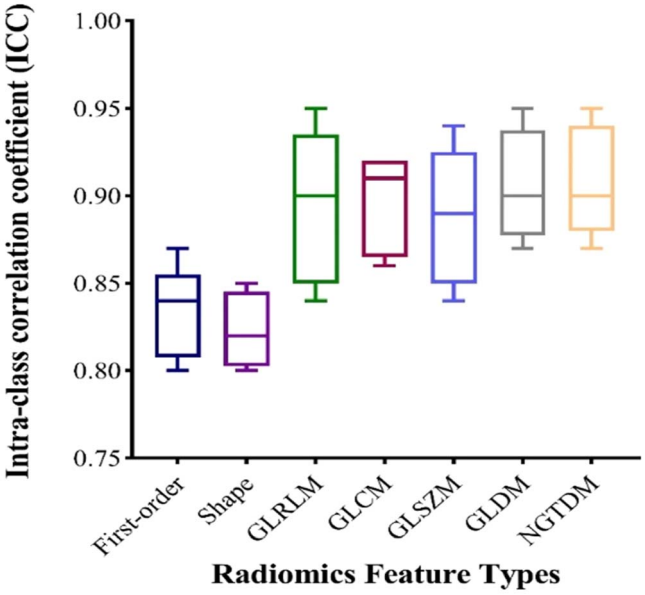

Figure 4. Intraclass correlation coefficients of the two observers for the different radiomics feature categories.

2.10. Statistics analyses

The statistical analyses were performed using SPSS 26 (Inc., Chicago, USA) and GraphPad Prism 8 software packages. The value of the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, precision, sensitivity, and specificity were calculated to evaluate the performance of the generated models using stratified 5-fold cross-validation.

3. Results

3.1. Patient characteristics and clinical outcomes

Patient demographic and clinical information is summarized in table 2. The patients were monitored and evaluated for up to 2 months after undergoing radiotherapy procedures.

Each patient had 13 dosimetric parameters extracted from the DVH. To facilitate comparison, figure 3 displays the means of the dosimetric parameters for the study participants in relation to proctitis and cystitis toxicities. Notably, patients with proctitis exhibited higher values in the dose parameters compared to those with cystitis.

3.2. Radiomics features

A total of 479 radiomics features, encompassing shape, first-order, and texture categories, were extracted from both original and filtered pretreatment planning CT and MRI T2W scans of the rectum and bladder walls. Out of these features, 59 were selected for analysis.

Figure 4 illustrates the ICC values for the various radiomics feature categories. The ICC values for the extracted features between the two physicians ranged from 0.8 to 0.95, signifying a strong agreement between the two observers. Notably, the ICC results for shape features were more influenced by the different image segmentations performed by the physicians, with an ICC value of 0.8.

3.3. Predictive radiomics analyses

The models were constructed using different sets of features: clinical features alone, CT and MRI radiomics features alone (as individual models), and a combination of clinical, dosimetric, and radiomics features (as the combined model), utilizing various machine learning techniques. The optimal methods and filters for each model are detailed in table 3. Predictive metrics for each model were computed using various feature selection methods applied to both original and filtered images of the rectum and bladder walls.

The clinical-only model, employing KNN and XGB classifiers with Chi2 and MRMR feature selections, exhibited the highest predictive performance for proctitis and cystitis, respectively. The pre-MRI T2W radiomics model for proctitis and cystitis prediction achieved its highest predictive performance using LR and LDA classifiers with LoG filter ($\sigma = 2$) and original images. MRMR and Anova feature selection algorithms displayed strong predictive performance for proctitis and cystitis, respectively.

In the case of the CT radiomics model, SVM and NB classifiers with LoG filter ($\sigma = 1$) and original images achieved the highest predictive performance for proctitis and cystitis, respectively. Lasso feature selection also demonstrated good predictive performance for both conditions.

Moreover, the combined MRI model, employing LDA and DT-RF classifiers with original images and LoG filters ($\sigma = 0.5 - 1.5$), and MRMR and RFE-Chi2

Table 3. Selected methods of individual and combined models for the proctitis and cystitis prediction.

		Clinical alone		Pre-MRI				Planning CT			
				Proctitis		Cystitis		Proctitis		Cystitis	
		Proctitis Individual model	Cystitis Individual model	Individual model	Combined model	Individual model	Combined model	Individual model	Combined model	Individual model	Combined model
Preprocessing	Filter	—	—	LOG ($\sigma = 2$)	Original	Original	LOG ($\sigma = 0.5, 1.5$)	LOG ($\sigma = 1$)	LOG ($\sigma = 2$)	Original	LOG ($\sigma = 0.5, 2$)
Models	Feature selection method	Chi2	MRMR	MRMR	MRMR	Anova	RFE-Chi2	Lasso	MRMR	Lasso	MRMR-Lasso
	Classifier	KNN	XGB	LR	LDA	LDA	DT-RF	SVM	XGB	NB	DT-XGB
Parameters	AUC (%)	68.33	64.27	71.66	91.0	69.07	96.57%	69.28	90.71	63.73	96.0
	Accuracy (%)	68.58	65.25	71.79	90.38	69.78	96.92	69.48	90.0	68.02	96.92
	Precision (%)	63.33	55.33	73.03	90.0	74.76	97.14	70.83	88.14	81.66	97.14
	Sensitivity (%)	70.47	64.76	70.0	93.33	54.0	96.50	63.33	93.0	40.0	96.0
	Specificity (%)	69.66	71.83	73.73	88.09	82.14	97.14	75.23	88.09	92.14	97.14

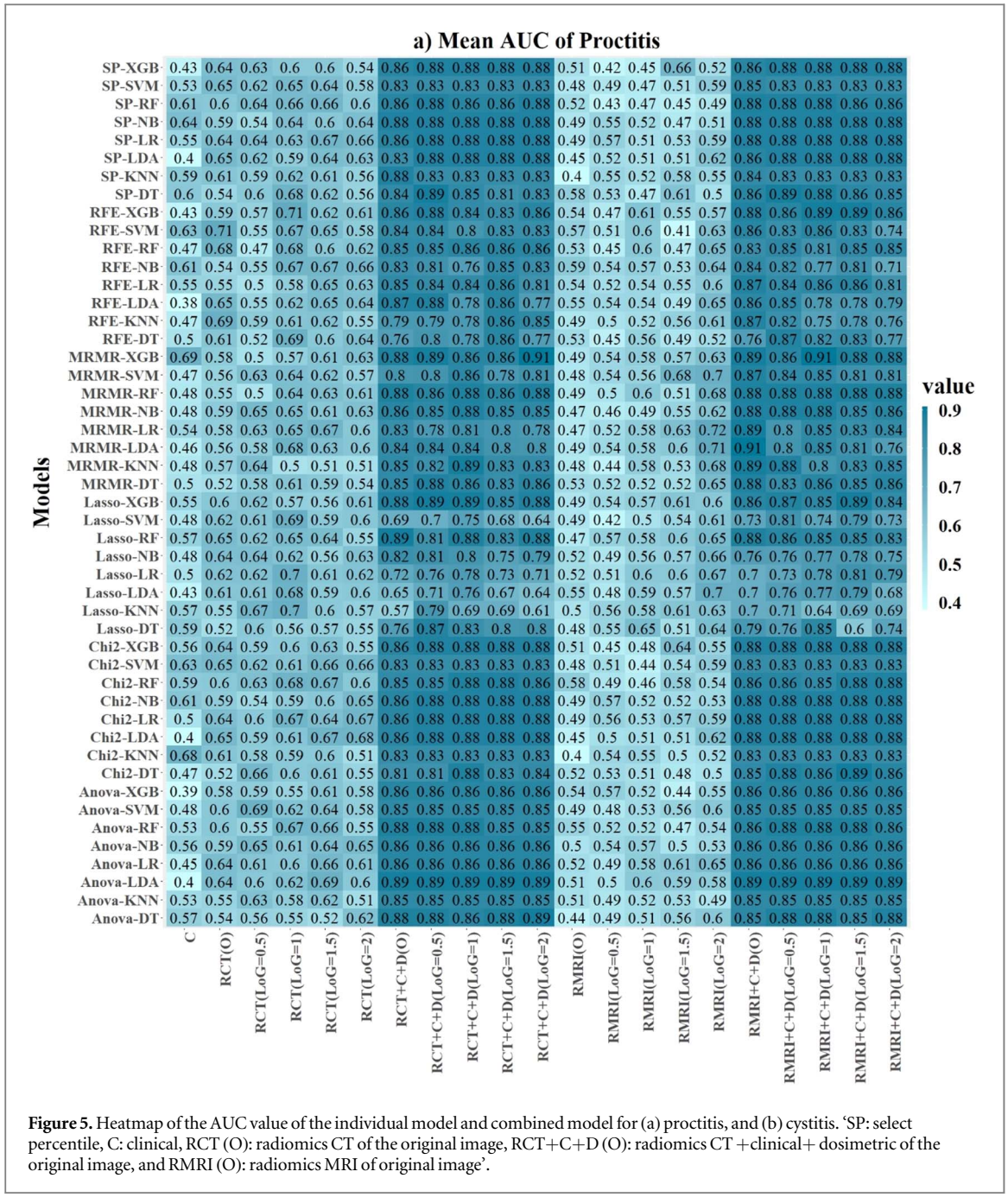

Figure 5. Heatmap of the AUC value of the individual model and combined model for (a) proctitis, and (b) cystitis. 'SP: select percentile, C: clinical, RCT (O): radiomics CT of the original image, RCT+C+D (O): radiomics CT +clinical+ dosimetric of the original image, and RMRI (O): radiomics MRI of original image'.

feature selections, delivered the highest predictive performance for proctitis and cystitis, respectively. Following closely were the combined CT model using XGB and DT-XGB classifiers with LoG filters ($\sigma = 2$) and LoG filters ($\sigma = 0.5 - 2$), alongside MRMR and MRMR-Lasso feature selections.

The results indicate that the combined model based on pretreatment planning CT or MRI outperforms the model based on planning CT or pre-MRI alone in terms of prediction ability.

The AUC values for the individual and combined models, employing various feature selection methods for proctitis and cystitis, are represented in a heatmap in figure 5. Detailed accuracy, precision, sensitivity, and specificity values for all models can be found in the Supplementary file (File 2).

3.4. Correlation of radiomics features and proctitis/cystitis toxicities

Out of the sixty-three rectal cancer patients in the study, here is the breakdown of the occurrence of proctitis, cystitis, and their combinations, as well as the number of patients without these toxicities based on the QLQ-C29 questionnaire:

Patients with proctitis (n = 48)

Patients with cystitis (n = 39)

Patients with both cystitis and proctitis (n = 38)

Patients with no toxicities (n = 25)

The popularity of the selected features for each method can be found in figures 6 and 7. For proctitis, MRI_FO_Min (n = 73), MRI_gldm_DEnt (n = 67), MRI_glszm_LGLZE (n = 87), MRI_glszm_Zvar (n = 83), MRI_Shape_Elong (n = 80), CT_ FO_Var

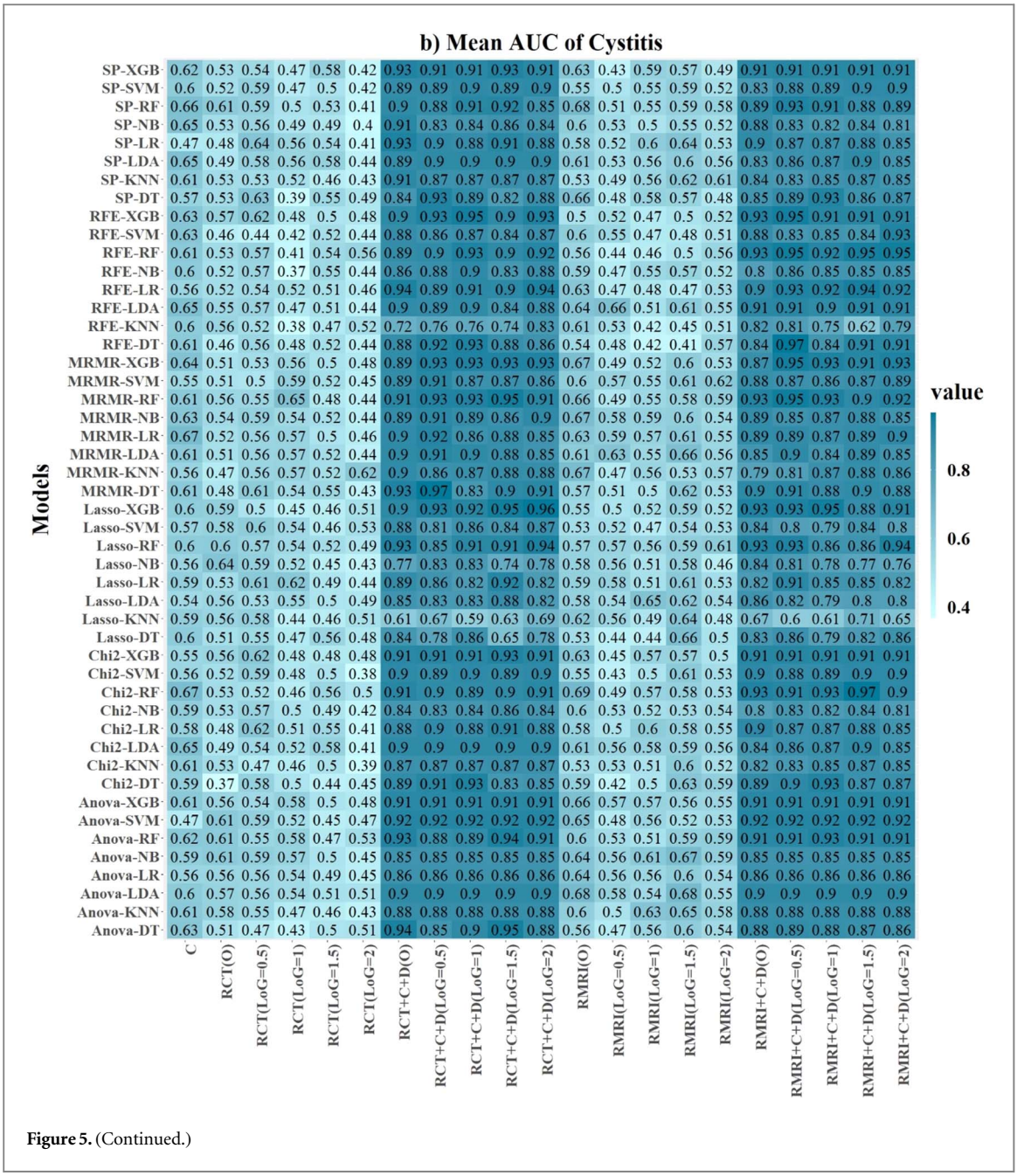

Figure 5. (Continued.)

(n = 79), CT_glcm_ClustPro (n = 127), CT_glcm_ClustSh (n = 133), CT_glcm_ClustTen (n = 88), CT_glcm_SumSq (n = 143), CT_gldm_GLV (n = 63), CT_glrlm_GLVar (n = 63), and CT_ngtdm_Contrast (n = 80) from radiomics features, D mean (n = 424), D40 (n = 291), D60 (n = 263), D70 (n = 327), D80 (n = 196), and D90 (n = 281) from dosimetric features, N stage (n = 89), and T stage (n = 82) from clinical features were selected frequently by the most of the six feature selection algorithms. In addition, the most significant features selected for cystitis were included MRI_FO_Min (n = 121), MRI_Shape_Elong (n = 72), MRI_Shape_MV (n = 80), MRI_Shape_SA (n = 85), MRI_ngtdm_Coars (n = 208), MRI_glrlm_RLNU (n = 160), MRI_gldm_DNU (n = 80), MRI_glszm_GLNU (n = 136), CT_glszm_LarAHGLE (n = 138), CT_glrlm_RLNU (n = 109), CT_glcm_Corr (n = 92), CT_glcm_Imc2 (n = 110), CT_gldm_DNU (n = 116), and CT_glszm_GLNU (n = 66) from radiomics features, D mean (n = 479), D min (n = 112), D50 (n = 130), D60 (n = 115), D70 (n = 235), D80 (n = 207), D90 (n = 269), and D100 (n = 192) from dosimetric features, T stage (n = 108), CEA (n = 65), N stage (n = 70), WallThick (n = 88), and TD (n = 100) from clinical features.

The full names and abbreviations of the features are provided in the Supplementary file (File 3).

4. Discussion

Medical imaging, as a clinical approach, plays a crucial role in assessing the side effects of the pathophysiological and functional processes associated with radiotherapy

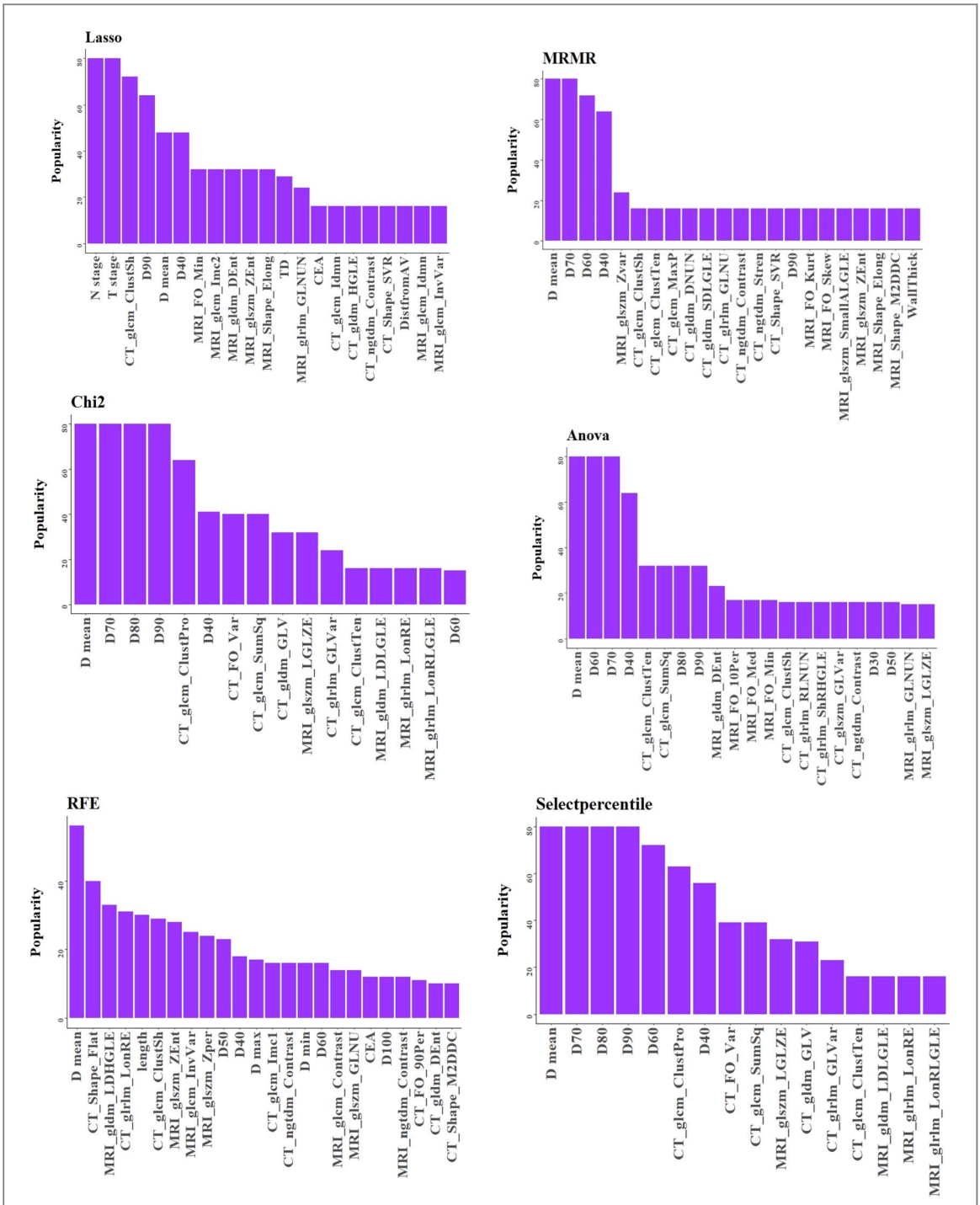

Figure 6. The most significant radiomics features that correlated with proctitis.

[15, 35, 36]. Therefore, identifying suitable predictive biomarkers for these complications is of great interest to clinicians and researchers. To the best of our knowledge, this study is the first to utilize radiomics features to predict acute rectal and bladder toxicities in locally advanced rectal cancer by combining different imaging modalities, namely CT and MRI.

Feature selection methods are employed to enhance the performance of predictive models and prevent overfitting. In this context, we utilized various classifiers and feature selection techniques to mitigate biases and uncertainties inherent to each method. The results revealed that the most effective feature selection and machine learning methods for the combined MRI model for proctitis and cystitis were MRMR/RFE-Chi2 and LDA/DT-RF classifiers with the original/LoG filter ($\sigma = 0.5/1.5$), respectively. Similarly, for the combined CT model, MRMR/MRMR-Lasso and XGB/DT-XGB with LoG filter ($\sigma = 2$)/LoG filter ($\sigma = 0.5/2$) yielded favorable results.

It's worth noting that previous studies have explored the effects of different feature selection methods and classifiers in the context of rectal cancer treatment response predictions. For instance, Shayesteh *et al* [28] investigated

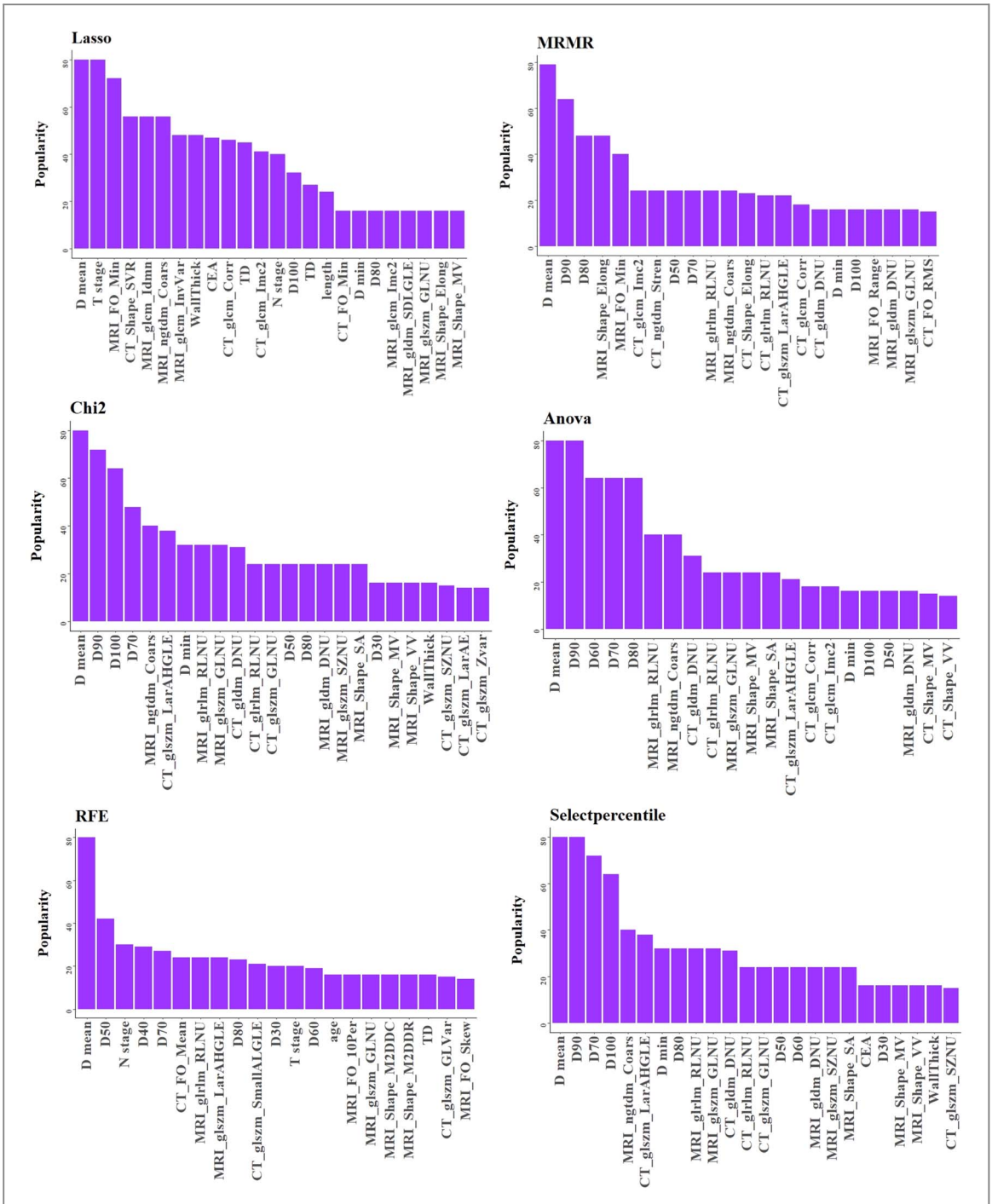

Figure 7. The most significant radiomics features that correlated with cystitis.

the impact of various feature selection methods on treatment response predictions for rectal cancer patients and found that feature selection algorithms improved machine learning performance, with the Cfs Subset Eval algorithm providing the best results. In another study [34], LR and SVM exhibited the highest model performances in predicting radiotherapy response in locally advanced rectal cancer patients using endorectal ultrasound images. The controversy surrounding the choice of different classifiers and feature selection methods suggests that there are no universally preferred classifiers and methods [37].

Figures 6 and 7 indicate that the top selected features, extracted by the six feature selection methods, include glcm_SumSq, glcm_ClustSh, glcm_ClustPro, glcm_ClustTen, glszm_LGLZE, and glszm_Zvar for the rectal wall. For the bladder wall, the most frequently selected features are ngtdm_Coars, glrlm_RLNU, glszm_LarAHGLE, glszm_GLNU, gldm_DNU, and glcm_Imc2. These findings align with a study by Mostafaei *et al* [4], which integrated CT-image features, clinical, and dosimetric parameters of prostate cancer to predict acute bladder and rectal injuries. They reported associations between radiomics features such as Small-dependence Low Gray-level Emphasis, High Gray-level Zone Emphasis, and Small-area Low Gray-level Emphasis with proctitis. Furthermore, Large-

dependence Low Gray-level Emphasis, Gray-level Variance, Short-run Low Gray-level Emphasis, and Small-area Low Gray-level Emphasis were related to cystitis. Another study by Hassaninejad *et al* [15] employed radiomic features extracted from rectal wall CT and MR images to predict radiation-induced rectal toxicity in prostate cancer patients. In their research, features including Coarseness, GrayLevelNonUniformity, Maximum2DdiameterColumn, SmallDependenceHighGrayLevelEmphasis, MCC, Contrast, and Flatness from the shape and texture families were identified as significant features associated with rectal toxicity. Based on the previous above-mentioned results and the current study, it can be concluded that texture features from various families, including GLSZM, GLCM, NGTDM, GLDM, and GLRLM, are predictive features for proctitis and cystitis. These features provide insights into the homogeneity or heterogeneity of tissues and can serve as valuable indicators for toxicity prediction. It's important to note that variations in radiomics features may be attributed to differences in the organs under study, image pre-processing techniques, and imaging modalities.

The results of our study indicate that combining radiomics features with clinical and dosimetric features can enhance the performance of the model. Specifically, the combined pre-MRI T2W model for predicting rectal and bladder toxicities demonstrated good performance, with mean AUC values of 0.85 and 0.87, respectively. Similarly, the combined CT model exhibited mean AUC values of 0.83 and 0.87 for predicting rectal and bladder toxicities. These findings align with a study by Bourbonne *et al* [38], who investigated radiomics features extracted from 3D dose maps and considered DVH and clinical parameters for predicting acute and late toxicities in both lungs and the esophagus. Their results showed that the combined model (clinical + DVH + radiomics) outperformed other models, with higher balanced accuracy, demonstrating the utility of incorporating radiomics features in predictive modeling.

Our study does have some limitations that could be addressed in future research. The small sample size of patients is a limitation, and future studies with larger patient cohorts are recommended to validate the results. Additionally, our study focused on predicting acute complications, and further research could explore the prediction of late effects. Lastly, considering radiomics features extracted from MR images obtained during or post-radiotherapy could potentially enhance the predictive ability of the models, offering a valuable avenue for future investigations.

5. Conclusion

In this study, we evaluated the potential of radiomics models to provide a quantitative and personalized assessment of radiation-induced toxicities in rectal cancer patients undergoing radiotherapy. We leveraged a diverse range of feature sources, including clinical data, dosimetric parameters, and radiomics features derived from both pretreatment CT and MR images. These features, serving as biomarkers, were employed to predict the development of proctitis and cystitis in these patients.

Our findings clearly indicate that the incorporation of clinical and dosimetric parameters significantly enhanced the predictive capabilities of our models for rectal and bladder toxicity following radiotherapy. This improvement was observed when compared to models based solely on clinical data, as well as those relying exclusively on pretreatment CT or MRI radiomics features. Notably, our results highlight the effectiveness of the LDA, DT, RF, and XGB classifiers in combination with feature selection algorithms such as MRMR, RFE, Chi2, and Lasso, especially when used in conjunction with the LoG filter. These models have the potential to aid in personalized decision-making for radiotherapy, offering valuable insights for patient management. However, it's essential to acknowledge that while these models show promise for clinical application, further refinement and validation are required. One critical aspect is the need for a larger training dataset to enhance the robustness and generalizability of these predictive models. This will be crucial for ensuring their effectiveness and reliability in real-world clinical settings.

Acknowledgments

This study was extracted from a PhD thesis from the first author approved by Tehran University of Medical Sciences. The authors would like to thank the staff of the Medical Imaging and Radiotherapy Oncology centers at Firouzgar Hospital of Iran University of Medical Sciences (Tehran, Iran).

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Funding

This work was supported by Tehran University of Medical Sciences, under Grant No. 49134.

Conflict of interest

The authors declare that they have no conflict of interest.

ORCID iDs

Samira Abbaspour https://orcid.org/0000-0001-6702-9172
Hamid Abdollahi https://orcid.org/0000-0003-0761-1309

Ghasem Hajainfar https://orcid.org/0000-0001-5359-2407
Mohammad Reza Ay https://orcid.org/0000-0002-5383-4418

References

[1] Rokan Z, Simillis C, Kontovounisios C, Moran B, Tekkis P and Brown G 2022 Locally recurrent rectal cancer according to a standardized MRI classification system: a systematic review of the literature *J. Clin. Med.* **11** 3511–27
[2] Moutabian H *et al* 2022 A systematic review of the therapeutic effects of resveratrol in combination with 5-fluorouracil during colorectal cancer treatment: with a special focus on the oxidant, apoptotic, and anti-inflammatory activities *Cancer Cell Int.* **22** 1–14
[3] Jaffray D A, Das S, Jacobs P M, Jeraj R and Lambin P 2018 How advances in imaging will affect precision radiation oncology *Int. J. Radiat. Oncol. Biol. Phys.* **101** 292–8
[4] Mostafaei S, Abdollahi H, Kazempour S, Isaac D and Abolfazl S 2020 CT imaging markers to improve radiation toxicity prediction in prostate cancer radiotherapy by stacking regression algorithm *Radiol Med.* **125** 87–97
[5] Appelt A L, Bentzen S M, Jakobsen A and Vogelius I R 2015 Dose-response of acute urinary toxicity of long-course preoperative chemoradiotherapy for rectal cancer *Acta. Oncol. (Madr).* **54** 179–86
[6] Luo Y *et al* 2018 A multiobjective bayesian networks approach for joint prediction of tumor local control and radiation pneumonitis in nonsmall-cell lung cancer (NSCLC) for response-adapted radiotherapy *Med. Phys.* **45** 3980–95
[7] Abdollahi H, Chin E, Clark H, Hyde D E, Thomas S, Wu J, Uribe C F and Rahmim A 2022 Radiomics-guided radiation therapy: opportunities and challenges *Phys. Med. Biol.* **67** 1–44
[8] Zhang W, Guo Y and Jin Q 2023 Radiomics and its feature selection: a review *Symmetry* **15** 1834
[9] Qin Y, Zhu L H, Zhao W, Wang J J and Wang H 2022 Review of radiomics-and dosiomics-based predicting models for rectal cancer *Front Oncol.* **12** 1–12
[10] Khajetash B, Talebi A, Bagherpour Z, Abbaspour S and Tavakoli M 2023 Introducing radiomics model to predict active plaque in multiple sclerosis patients using magnetic resonance images *Biomed. Phys. Eng. Express* **9** 55004
[11] Guzzo A, Fortino G, Greco G and Maggiolini M 2023 Data and model aggregation for radiomics applications: emerging trend and open challenges *Inf. Fusion* **100** 101923
[12] Di Costanzo G *et al* 2023 Artificial intelligence and radiomics in magnetic resonance imaging of rectal cancer: a review *Explor Target Anti-tumor Ther.* **4** 406–21
[13] Ger R B, Wei L, Naqa I E and Wang J 2023 The promise and future of radiomics for personalized radiotherapy dosing and adaptation *Semin Radiat Oncol.* **33** 252–61
[14] Lucia F *et al* 2021 Radiomics analysis of 3d dose distributions to predict toxicity of radiotherapy for cervical cancer *J. Pers. Med.* **11** 1–12
[15] Hassaninejad H, Abdollahi H, Abedi I, Amouheidari A and Tavakoli M B 2023 Radiomics based predictive modeling of rectal toxicity in prostate cancer patients undergoing radiotherapy: CT and MRI comparison *Phys. Eng. Sci. Med.* 1–18
[16] Yang Z *et al* 2023 Machine-learning with region-level radiomic and dosimetric features for predicting radiotherapy-induced rectal toxicities in prostate cancer patients *Radiother Oncol* **183** 109593
[17] Abdollahi H *et al* 2019 MRI radiomic analysis of IMRT-induced bladder wall changes in prostate cancer patients: a relationship with radiation dose and toxicity *J. Med. Imaging Radiat. Sci.* **50** 252–60
[18] Abdollahi H *et al* 2018 Rectal wall MRI radiomics in prostate cancer patients: prediction of and correlation with early rectal toxicity *Int. J. Radiat Biol.* **94** 829–37
[19] Berger T *et al* 2023 Assessing the generalisability of radiomics features previously identified as predictive of radiation-induced sticky saliva and xerostomia *Phys. Imaging Radiat. Oncol.* **25** 100404
[20] Berger T *et al* 2023 Sub-regional analysis of the parotid glands: model development for predicting late xerostomia with radiomics features in head and neck cancer patients *Acta Oncol.* **62** 166–73
[21] Feng B *et al* 2023 Pseudo-siamese network combined with dosimetric and clinical factors, radiomics features, CT images and 3D dose distribution for the prediction of radiation pneumonitis: a feasibility study *Clin. Transl. Radiat. Oncol.* **38** 188–94
[22] Qin Q *et al* 2020 Cone-beam CT radiomics features might improve the prediction of lung toxicity after SBRT in stage I NSCLC patients *Thorac Cancer* **11** 964–72
[23] Chen J *et al* 2018 Investigating rectal toxicity associated dosimetric features with deformable accumulated rectal surface dose maps for cervical cancer radiotherapy *Radiat. Oncol.* **13** 1–10
[24] Mavroidis P *et al* 2018 Fitting NTCP models to bladder doses and acute urinary symptoms during post- prostatectomy radiotherapy *Radiat. Oncol.* **13** 1–8
[25] Ribault A *et al* 2023 Circulating microvesicles correlate with radiation proctitis complication after radiotherapy *Sci. Rep.* **13** 1–12
[26] Liu Z *et al* 2017 Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer *Clin. Cancer Res.* **23** 7253–62
[27] Whistance R N *et al* 2009 Clinical and psychometric validation of the EORTC QLQ-CR29 questionnaire module to assess health-related quality of life in patients with colorectal cancer *Eur. J. Cancer* **45** 3017–26
[28] Shayesteh S P *et al* 2020 Author correction: prediction of response to neoadjuvant chemoradiotherapy by MRI-Based machine learning texture analysis in rectal cancer patients *J. Gastrointest Cancer.* **51** 610
[29] Shayesteh S P *et al* 2019 Neo-adjuvant chemoradiotherapy response prediction using MRI based ensemble learning method in rectal cancer patients *Phys. Medica.* **62** 111–9
[30] Shiri *et al* 2020 Repeatability of radiomic features in magnetic resonance imaging of glioblastoma: test–retest and image registration analyses *Med. Phys.* **47** 4265–80
[31] Van Griethuysen J J M *et al* 2017 Computational radiomics system to decode the radiographic phenotype *Cancer Res.* **77** e104–7
[32] Berrar D 2019 Cross-Validation *Encyclopedia of Bioinformatics and Computational Biology* (Oxford: Academic) 542–5
[33] Szeghalmy S and Fazekas A 2023 A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning *Sensors* **23** 2333–60
[34] Abbaspour S *et al* 2022 Endorectal ultrasound radiomics in locally advanced rectal cancer patients: despeckling and radiotherapy response prediction using machine learning *Abdom. Radiol.* **2022** 1–15
[35] Ritlumlert N *et al* 2023 Improved prediction of radiation - induced hypothyroidism in nasopharyngeal carcinoma using pre - treatment CT radiomics *Sci. Rep.* **13** 17437
[36] Smyczynska U *et al* 2021 Prediction of radiation-induced hypothyroidism using radiomic data analysis does not show superiority over standard normal tissue complication models *Cancers (Basel)* **13** 5584–603
[37] Wang J *et al* 2019 Radiomics features on radiotherapy treatment planning CT can predict patient survival in locally advanced rectal cancer patients *Sci. Rep.* **9** 1–9
[38] Visvikis D, Hatt M and Schick U 2020 Radiomics analysis of 3D dose distributions to predict toxicity of radiotherapy for lung cancer *Radiother Oncol.* **155** 144–50