European Radiology (2019) 29:6182–6190
https://doi.org/10.1007/s00330-019-06222-8

IMAGING INFORMATICS AND ARTIFICIAL INTELLIGENCE

Radiomics analysis of multiparametric MRI for the preoperative evaluation of pathological grade in bladder cancer tumors

Huanjun Wang[1] · Daokun Hu[2,3] · Haohua Yao[4] · Maodong Chen[2,3] · Shurong Li[1] · Haolin Chen[2,3] · Junhang Luo[4] · Yanqiu Feng[2,3] · Yan Guo[1]

Received: 26 February 2019 / Revised: 2 April 2019 / Accepted: 3 April 2019 / Published online: 23 April 2019

Abstract

Objectives To develop and validate an MRI-based radiomics strategy for the preoperative estimation of pathological grade in bladder cancer (BCa) tumors.

Methods A primary cohort of 70 patients (31 high-grade BCa and 39 low-grade BCa) with BCa were retrospectively enrolled. Three sets of radiomics features were separately extracted from tumor volumes on T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and apparent diffusion coefficient (ADC) maps. Two sets of multimodal features were separately generated by the maxout and concatenation of the above mentioned single-modality features. Each feature set was subjected to a two-sample *t* test and the least absolute shrinkage and selection operator (LASSO) algorithm for feature selection. Multivariable logistic regression (LR) analysis was used to obtain five corresponding radiomics models. The diagnostic abilities of the radiomics models were evaluated using receiver operating characteristic (ROC) curve analysis and compared using the DeLong test. Validation was performed on a time-independent cohort containing 30 consecutive patients.

Results The areas under the ROC curves (AUCs) of single-modality T2WI, DWI, and ADC models in the training cohort were 0.7933 (95% confidence interval [CI] 0.7471–0.8396), 0.8083 (95% CI 0.7565–0.8601), and 0.8350 (95% CI 0.7924–0.8776), respectively. Both multimodality models achieved higher AUCs (maxout 0.9233, 95% CI 0.9001–0.9466; concatenation 0.9233, 95% CI 0.9001–0.9466) than single-modality models. The AUCs of the maxout and concatenation models in the validation cohort were 0.9186 and 0.9276, respectively.

Conclusions The MRI-based multiparametric radiomics approach has the potential to be used as a noninvasive imaging tool for preoperative grading of BCa tumors. Multicenter validation is needed to acquire high-level evidence for its clinical application.

Key Points

- *Multiparametric MRI may help in the preoperative grading of BCa tumors.*
- *The Joint_Model established from T2WI, DWI, and ADC feature subsets demonstrated a high diagnostic accuracy for preoperative prediction of pathological grade in BCa tumors.*
- *The radiomics approach has the potential to preoperatively assess tumor grades in BCa and avoid subjectivity.*

Keywords Magnetic resonance imaging · Urinary bladder · ROC curve · Regression analysis

Huanjun Wang and Daokun Hu contributed equally to this work.

✉ Yanqiu Feng
foree@fimmu.com

✉ Yan Guo
dr.guoyan@163.com

[1] Department of Radiology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, People's Republic of China

[2] School of Biomedical Engineering, Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou, Guangdong, People's Republic of China

[3] Key Laboratory of Mental Health of the Ministry of Education, School of Biomedical Engineering, Southern Medical University, Guangzhou, Guangdong, People's Republic of China

[4] Department of Urology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, People's Republic of China

Abbreviations

ADC	Apparent diffusion coefficient
AUC	Area under the curve
BCa	Bladder cancer
CI	Confidence interval
DTI	Diffusion tensor imaging
DWI	Diffusion-weighted imaging
LASSO	Least absolute shrinkage and selection operator
MIBCa	Muscle-invasive bladder cancer
NMIBCa	Non-muscle-invasive bladder cancer
PWI	Perfusion-weighted imaging
ROC	Receiver operating characteristic
TIS	Tumor in situ

Introduction

Bladder cancer (BCa) ranks as the ninth most common cancer worldwide and is one of the most expensive human malignancies to manage [1–3]. BCa can be histologically stratified into low- and high-grade cancers [4]. Although most non-muscle-invasive bladder cancers (NMIBCas) are low grade and often have an indolent natural history [5–8], one-third of NMIBCas consist of high-grade BCas, which can progress to muscle invasion and metastases in approximately 20–25% of patients [5–9].

Management of BCa is based on the pathological findings from a biopsy or transurethral resection, with particular attention to histology, grade, and depth of invasion. Presentations with high-risk features, e.g., multiple high-grade T1 tumors with tumor in situ (TIS) or increased depth of invasion, may be considered for cystectomy [10]. Thus, transurethral resection or cystoscopic biopsy remains the standard for the preoperative staging and grading of BCa. However, transurethral resection of BCa tumors carries a significant risk of understaging the cancer, and biopsy results are not always representative of the entire tumor and may require a repeated biopsy. The recently published study regarding VI-RADS based on multiparametric MRI for the discrimination of NMIBCa from muscle-invasive bladder cancer (MIBCa) showed that the recent proposed VI-RADS is accurate in differentiating MIBCa from NMIBCa and further suggests that multiparametric MRI is the best imaging modality for local staging [11]. On the other hand, an accurate preoperative prediction method for tumor grade would also be crucial for personal management.

Our previous studies [12] have shown that with an apparent diffusion coefficient (ADC) cutoff value of 0.899×10^{-3} mm^2/s, the ADC value of low-grade tumors is significantly higher than that of high-grade tumors. Nevertheless, as it is the only predictive index, a single ADC value is not strong evidence and yields insufficient information to predict the differentiation ability.

In recent years, the "radiomics approach" has been increasingly applied, and with the development of high-throughput technology and analytical approaches [13–16], radiomics data may aid in disease detection, diagnosis, evaluation of prognosis, and prediction of treatment response [15]. Radiomics helps to make a diagnosis while avoiding invasive procedures. In a recent small sample size study [17], radiomics analysis showed that textural features from diffusion-weighted imaging (DWI) and ADC maps can reflect the difference between low- and high-grade BCas and can preoperatively facilitate image-based BCa grading, but the information provided only by a single DWI or ADC value was limited. In addition, an independent validation set was not evaluated.

According to the European Organisation for Research and Treatment of Cancer (EORTC) risk table, low-risk BCa tumors can be defined according to the tumor number, size, prior recurrence rate, T category, carcinoma in situ, and tumor grade, which can be assessed through clinical information and radiomics data analysis. Patients with low-risk tumors can choose a surveillance option in clinics instead of an invasive biopsy procedure; however, whether all of these are feasible by using the radiomics approach also remains unknown.

Hence, we conducted this study aiming to develop and validate a radiomics approach based on multiparametric MRI signatures for preoperative evaluation of the histological grade of BCa tumors.

Patients and methods

Patients

Ethical approval was obtained for this retrospective study, and informed consent from patients was waived. Patients who underwent transurethral resection or radical cystectomy and had tumor grades determined were selected by evaluation of the institution's database of medical records from February 2013 to October 2017. A total of 100 BCa patients with 100 tumors were ultimately included, and only the largest tumor detected on MRI was selected in patients with multiple tumors. The tumor dimension ranged from 12 to 75 mm, with a mean ± standard deviation (SD) of 35.08 ± 15.16 mm. Among the 100 tumors, the tumor stage was characterized as TIS, Ta, T1, T2, and T3 in 1, 9, 68, 17, and 5 cases, respectively. Patients were then divided into two independent sets based on the time of surgery; the 70 patients who were treated earlier constituted the training set, and the remaining 30 patients constituted the validation set.

MRI data acquisition

Data from all patients who underwent bladder MRI, including T2WI and DWI, were included. Pelvic MRI was performed

on all patients in the supine position after a 4–6-h fast using a 3-T MRI system (Magnetom Trio, Siemens Healthineers) with an eight-channel phased-array pelvic coil. Axial fast spin-echo T2WI (3500–4500 ms/91–131 ms) and high–spatial resolution (320 × 320 matrix) fat-suppressed T2WI (3500–4650 ms/95–102 ms) in three orthogonal planes were acquired with three excitations, a field of view of 220 mm × 220 mm, and a slice thickness of 4 mm. DWI was performed with free-breathing, water-excited, single-shot, spin-echo, echo-planar imaging (4000 ms/78 ms) using b values of 0 and 1000 s/mm^2 and eight excitations in the axial and oblique sagittal planes perpendicular to the tumor base [18]. All images were obtained with a 4-mm slice thickness and a 0.4-mm intersection gap. ADC maps were calculated voxel by voxel with the monoexponential model using formula (1):

$$S_{1000}/S_0 = \exp\left(-b_{1000} \times \mathrm{ADC}\right) \quad (1)$$

where S_{1000} and S_0 are the signal intensities with and without a diffusion gradient, respectively. T2WI and DWI sequences were retrieved for image feature extraction.

ROI delineation

MR images of all patients were analyzed by two radiologists (W.H.J., with 8 years of experience in BCa imaging, and L.S.R., with 10 years of experience in BCa imaging); both were blinded to the histopathology results. Region of interest (ROI) segmentation based on T2WI and DWI sequences was independently and manually performed slice by slice through the whole tumor, excluding areas of necrosis or vessels. On T2WI sequences, ROIs were drawn along the contour of the tumor as visually defined by slightly high signal in the bladder lumen. On DWI (b value of 1000 s/ mm^2) sequences, ROIs were placed in hyperintense areas without susceptibility artifacts. Divergence was solved by discussion until a consensus was reached.

Feature extraction

MRI data for each patient, other than ADC maps, were normalized to minimize the effects of brightness and contrast variations on radiomics feature extraction. Specifically, for each volume containing the lesion, intensity values at the 1st and 99th percentiles were computed and used to scale the intensity value to [0, 1] through the min-max method and then rebinned to 8 bits/pixel.

For each tumor area, all radiomics texture features were calculated through coordinates on the X-Y plane and averaged for all slices within the three-dimensional tumor volume using the open-source Pyradiomics package (https://github.com/Radiomics/pyradiomics) [19]. A total of 924 features were extracted from a single MR image, including 14 shape-based features, 220 gray-level cooccurrence matrix (GLCM) features, 160 gray-level run length matrix (GLRLM) features, 160 gray-level size zone matrix (GLSZM) features, 50 neighborhood gray tone difference matrix (NGTDM) features, 140 neighboring gray-level dependence matrix (GLDM) features, and 180 first-order statistics features (Table 1). To evaluate the performance of radiomics features from single and multimodality images, five feature subsets were constructed as follows. First, three separate feature subsets were generated from T2WI, DWI, and ADC maps. Second, the above three feature subsets were combined to establish a joint feature subset. Finally, for every single feature, three feature values were obtained from the above three subsets, and the maximum value was selected to construct the new maxout subset. The raw texture feature values from the five radiomics feature subsets were then exported for further classification and statistical analysis. The process of radiomics analysis is shown in Fig. 1.

Feature selection and classification

To identify the most discriminating radiomics feature set in the training cohort, two feature selection procedures were applied on the previously mentioned five feature subsets. First, a two-sample t test was performed on each feature subset to select the best features that were significantly different between the low-grade and the high-grade groups. Second, the least absolute shrinkage and selection operator (LASSO) feature selection analysis [20] was applied before classification (Fig. 2). Let Z represent the linear combination of the selected features and their corresponding LASSO regression coefficients, and the ultimate radiomics score (Rad-score) of each patient can be obtained by formula (2):

$$\text{Rad-score} = 1/(1 + \exp\left(-Z\right)) \quad (2)$$

The performance metric was assessed using the area under the receiver operating characteristic (ROC) curve (AUC) and the Youden index [21, 22]. To avoid feature selection bias and overfitting, we used a tenfold (onefold held out for testing), patient-stratified, cross-validation scheme, which was repeated 50 times.

For each feature subset, based on the characteristics selected in the above procedures, a logistic regression (LR) classifier was first trained on a training cohort and evaluated with the independent validation cohort. The classification performance of all five models was then evaluated using accuracy, sensitivity, specificity, and AUC values. Specifically, the performance evaluation of the training cohort was used as the average across 50 runs of tenfold cross-validation. The diagnostic abilities of the five models were finally compared using the DeLong test. Figure 3 shows the workflow of this study.

Table 1 The list of all the features used in this study. In our study, except for the features in the first feature group that were calculated using one parameter, features in the remaining 6 feature groups were calculated using 10 parameters. In the end, we obtained 924 features from a single MR image

Feature classes	Number of features	Definite features
Shape-based	14	SphericalDisproportion, Maximum2DDiameterSlice, Sphericity, MinorAxis, Elongation, SurfaceVolumeRatio, Volume, Maximum3DDiameter, MajorAxis, SurfaceArea, Flatness, LeastAxis, Maximum2DDiameterColumn, Maximum2DDiameterRow
Gray-level cooccurrence Matrix (GLCM)	22	JointAverage, JointEntropy, ClusterShade, MaximumProbability, Idmn, JointEnergy, Contrast, DifferenceEntropy, InverseVariance, DifferenceVariance, Idn, Idm, Correlation, Autocorrelation, SumEntropy, SumSquares, ClusterProminence, Imc2, Imc1, DifferenceAverage, Id, ClusterTendency
Gray-level run length matrix (GLRLM)	16	ShortRunLowGrayLevelEmphasis, GrayLevelVariance, LowGrayLevelRunEmphasis, RunVariance, GrayLevelNonUniformityNormalized, GrayLevelNonUniformity, LongRunEmphasis, ShortRunHighGrayLevelEmphasis, RunLengthNonUniformity, ShortRunEmphasis, LongRunHighGrayLevelEmphasis, RunPercentage, LongRunLowGrayLevelEmphasis, RunEntropy, HighGrayLevelRunEmphasis, RunLengthNonUniformityNormalized
Gray-level size zone matrix (GLSZM)	16	GrayLevelVariance, SmallAreaLowGrayLevelEmphasis, GrayLevelNonUniformityNormalized, SizeZoneNonUniformityNormalized, SizeZoneNonUniformity, GrayLevelNonUniformity, LargeAreaEmphasis, SmallAreaHighGrayLevelEmphasis, ZonePercentage, ZoneEntropy, LargeAreaLowGrayLevelEmphasis, LargeAreaHighGrayLevelEmphasis, ZoneVariance, HighGrayLevelZoneEmphasis, SmallAreaEmphasis, LowGrayLevelZoneEmphasis,
Neighborhood gray tone difference matrix (NGTDM)	5	Coarseness, Complexity, Strength, Contrast, Busyness
Neighborhood gray-level dependence matrix (GLDM)	14	GrayLevelVariance, HighGrayLevelEmphasis, DependenceEntropy, DependenceNonUniformity, GrayLevelNonUniformity, SmallDependenceEmphasis, SmallDependenceHighGrayLevelEmphasis, DependenceNonUniformityNormalized, DependenceVariance, LargeDependenceEmphasis, LargeDependenceLowGrayLevelEmphasis, LargeDependenceHighGrayLevelEmphasis, SmallDependenceLowGrayLevelEmphasis, LowGrayLevelEmphasis,
First-order statistic	18	InterquartileRange, Skewness, Uniformity, Median, Energy, RobustMeanAbsoluteDeviation, MeanAbsoluteDeviation, TotalEnergy, Maximum, RootMeanSquared, 90Percentile, Minimum, Entropy, Range, Variance, 10Percentile, Kurtosis, Mean

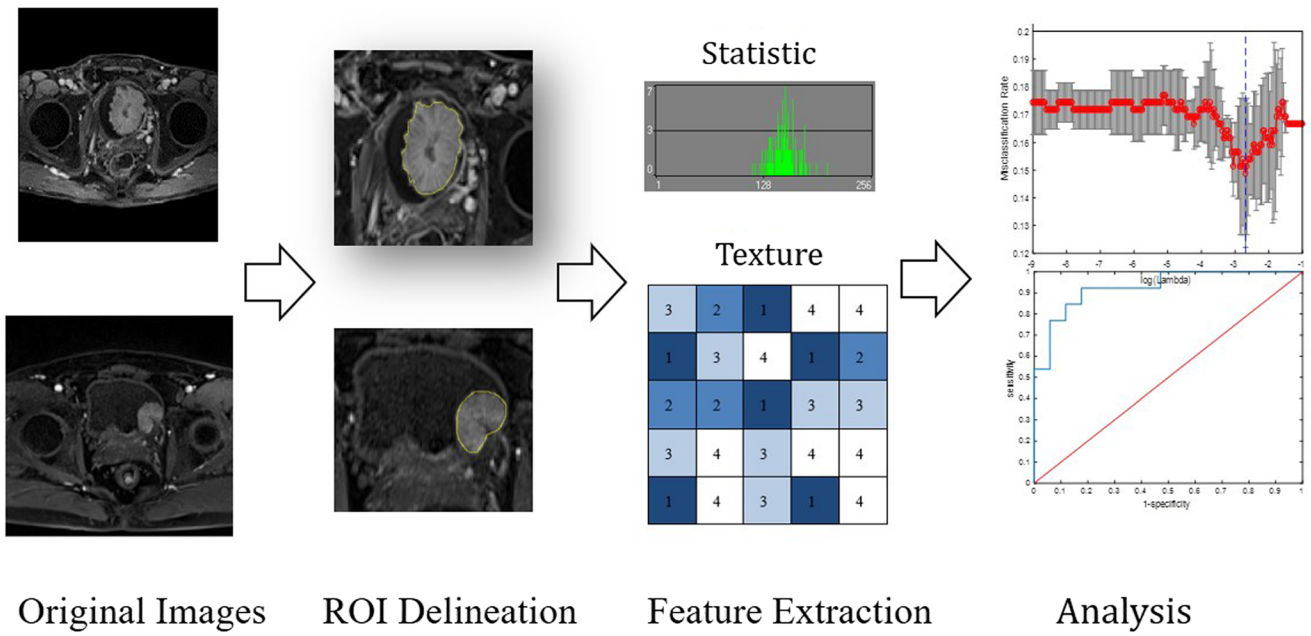

Fig. 1 Radiomics analysis workflow. From high-quality medical images, an ROI was defined via automated, semiautomated, or manual segmentation methods. Statistical and quantitative features were extracted from the ROI. These features were then subjected to analytic procedures. Feature selection and model construction steps were usually included in the analytic procedures

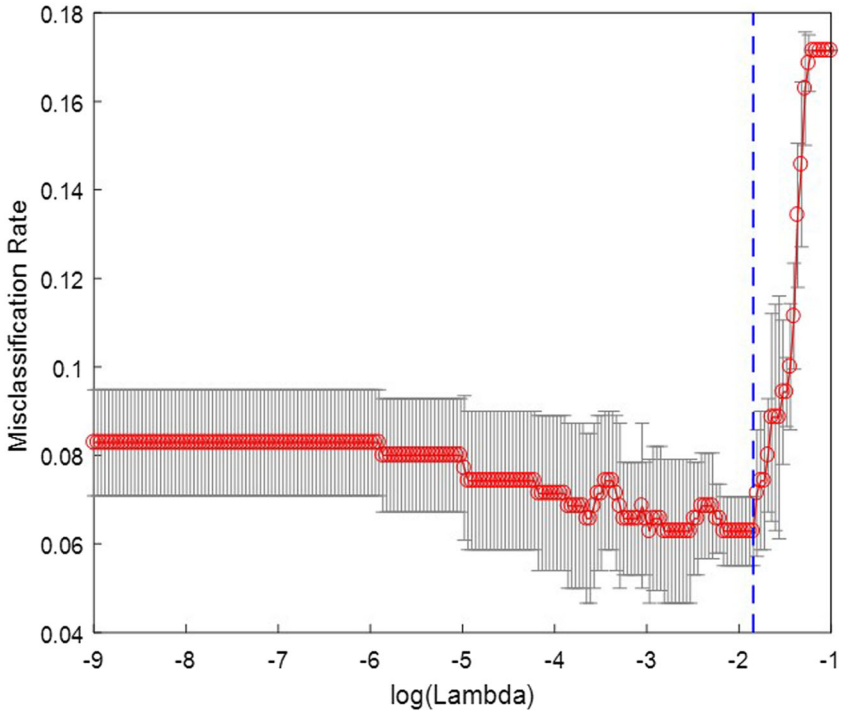

Fig. 2 Feature selection using the LASSO regression method according to 10-fold cross-validation. On the basis of minimum criteria, we selected tuning parameters (λ). The vertical blue line indicates the optimal value of λ, where the minimum misclassification rate from the LASSO regression cross-validation procedure was obtained. According to 10-fold cross-validation, an optimal λ value of 0.159 with log (λ) = − 1.839 was chosen

Statistical analysis

Statistical analysis of the demographics and tumor grades in the training and validation sets was assessed using SPSS software (version 22, IBM). All of the other analysis processes mentioned above were implemented in MATLAB R2017A (MathWorks). In all cases, p values less than 0.05 were considered statistically significant.

Results

In total, 100 BCa patients (86 men, 14 women; age range 40–86 years, mean 63.4 ± 10.4 years) with 100 tumors were finally included. Based on the time of surgery, the 70 patients who were treated earlier were grouped into the training cohort, and the remaining 30 patients were included in the validation cohort. Table 2 shows the demographics and tumor grades in the two cohorts. High-grade BCa constituted 46.99% (31/70) and 46.88% (13/30) of the training and validation sets, respectively, and low-grade BCa constituted 53.01% (39/70) and 53.12% (17/30) of the training and validation sets, respectively. No significant difference was found between the two groups ($p = 0.931$, t test).

The feature numbers selected from each original feature subset through the aforementioned two-step feature selection strategies are shown in Table 3. The estimated performance of the five radiomics models on the primary and validation cohorts is shown in Table 4. The results showed that the Joint_Model performed best (AUC = 0. 9276) in the validation cohort, compared with the respective AUCs of 0.8235, 0.7692, 0.8824, and 0.9186 in the remaining four models. The trends in sensitivity and specificity of the five models were the same as those in the AUC values (Table 4). Through the DeLong test, we found significant difference between the Joint_Model and the T2WI_Model ($p < 0.05$) and between the DWI_Model ($p < 0.05$) and the ADC_Model ($p = 0.063$), while no significant difference was found between the Joint_Model and the Max-out_Model ($p = 0.817$). The ROC curves of the five radiomics models in the validation set are shown in Fig. 4. In conclusion, the performance of the Joint_Model exceeded that of the other four models.

In the Joint_Model, 2772 features were compressed into 7 features (including 1 feature from the T2WI feature subset, 1 feature from the DWI feature subset, and 5 features from the ADC feature subset) with nonzero coefficients as determined by the LASSO algorithm in the training cohort (Fig. 2). These selected features included the following:

ADC_ log-sigma-5-0-mm-3D_firstorder_Mean,
ADC_ log-sigma-2-0-mm-3D_firstorder_Kurtosis,
ADC_ original_glcm_Imc2,
ADC_ wavelet-LL_firstorder_Minimum,
DWI_ log-sigma-4-0-mm-3D_glcm_Imc1,
T2WI_ wavelet-LL_firstorder_Mean, and
ADC_ wavelet-LL_firstorder_10Percentile.

The linear combination of the features and their corresponding coefficients was as follows: $Z = -0.4752 + \text{ADC\_log-sigma-5-0-mm-3D\_firstorder\_Mean} \times 1.4857 + \text{ADC\_log-sigma-2-0-mm-3D\_firstorder\_Kurtosis} \times 1.8191 + \text{ADC\_original\_glcm\_Imc2} \times 0.7614 + \text{ADC\_ wavelet-}$

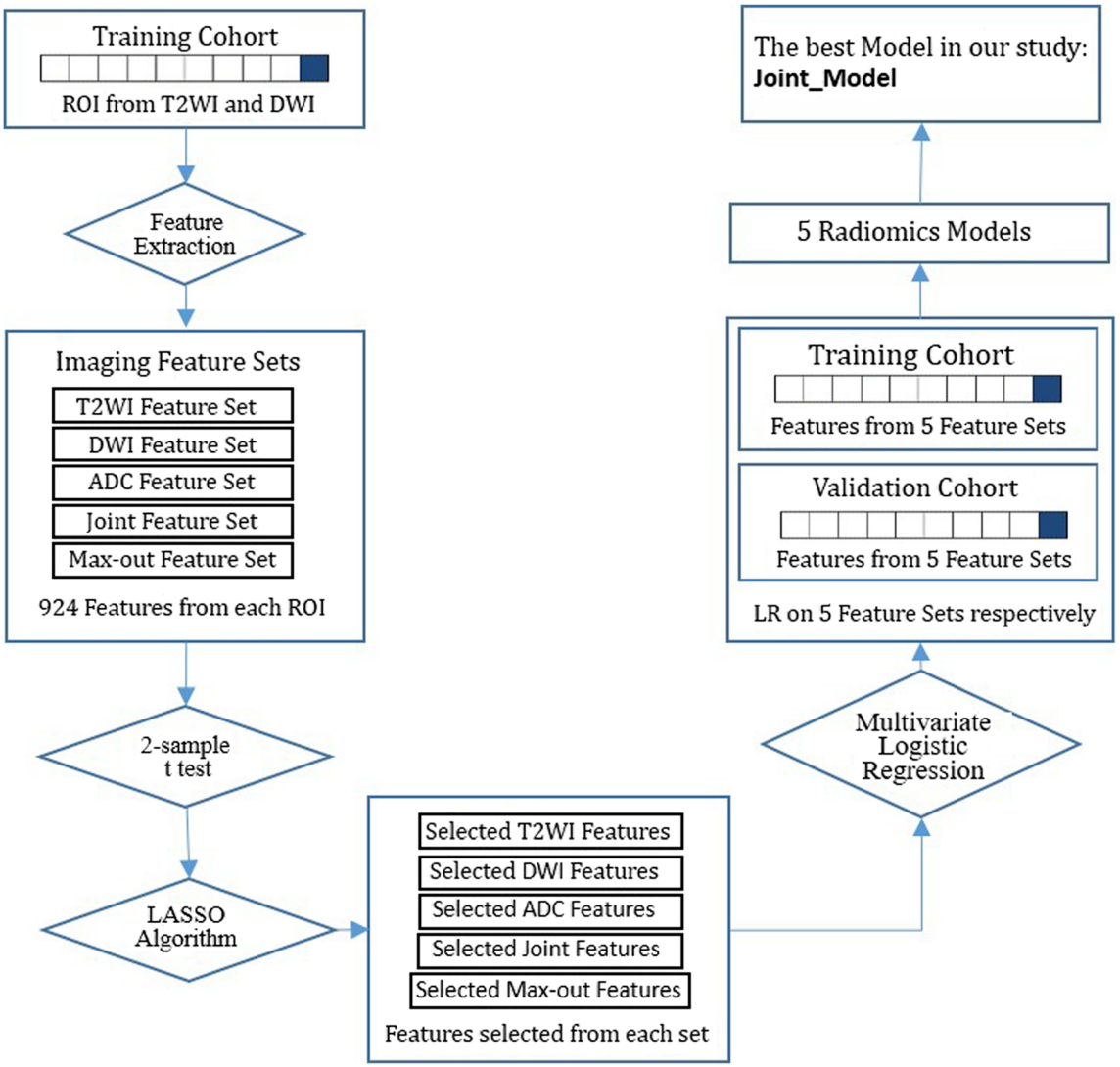

Fig. 3 Workflow of this study

LL_firstorder_Minimum × 0.4611 + DWI_ log-sigma-4-0-mm-3D_glcm_Imc1 × − 1.3684 + T2WI_ wavelet-LL_firstorder_Mean × − 0.7790 + ADC_ wavelet-LL_firstorder_10Percentile × − 1.1853.

Table 2 Demographic, tumor grade, and stage data in the training and validation sets

	Training set	Validation set	*p* value
Age (mean ± SD)	62.6 ± 11.0	65.2 ± 8.1	0.175
Gender (No. (%))			0.456
Male	59 (85.54)	27 (90.63)	
Female	11 (14.46)	3 (9.37)	
Grade (No. (%))			0.931
High grade	31 (46.99)	13 (46.88)	
Low grade	39 (53.01)	17 (53.12)	
Pathologic stage (No. (%))			–
≤T1	53 (75.71)	25 (83.33)	
≥T2	17 (24.29)	5 (16.67)	

Discussion

In our study, we developed and validated a multiparametric MRI-based radiomics approach for preoperative, noninvasive

Table 3 Feature numbers selected from each original feature subset through the following 2-step feature selection strategies: the 2-sample *t* test ($p < 0.01$) was applied in step 1, and the LASSO algorithm was applied in step 2

	T2WI	DWI	ADC	Max_out	Joint
Original	924	924	924	924	2772
Selected features after step 1	228	235	360	280	823
Selected features after step 2	21	14	14	5	7

Table 4 Classification performance of the 5 radiomics models in the training and validation cohorts

		Accuracy	Sensitivity	Specificity	AUC
T2WI_Model	Training cohort (95% CI)	0.7743 (0.7311–0.8174)	0.7467 (0.6764–0.8170)	0.7967 (0.7438–0.8495)	0.7933 (0.7471–0.8396)
	Validation cohort	0.7667	0.7692	0.7647	0.7828
DWI_Model	Training cohort (95% CI)	0.7257 (0.6710–0.7805)	0.7067 (0.6321–0.7812)	0.7400 (0.6847–0.7953)	0.8083 (0.7565–0.8601)
	Validation cohort	0.7667	0.7692	0.7647	0.7692
ADC_Model	Training cohort (95% CI)	0.8171 (0.7779–0.8564)	0.8100 (0.7335–0.8865)	0.8250 (0.7767–0.8733)	0.8350 (0.7924–0.8776)
	Validation cohort	0.8000	0.8462	0.7647	0.8054
Max-out_Model	Training cohort (95% CI)	0.8429 (0.7967–0.8890)	0.8283 (0.7533–0.9034)	0.8533 (0.8004–0.9062)	0.8850 (0.8413–0.9287)
	Validation cohort	0.8333	0.7692	0.8824	0.9186
Joint_Model	Training cohort (95% CI)	0.8543 (0.8181–0.8905)	0.8267 (0.7609–0.8925)	0.8783 (0.8330–0.9236)	0.9233 (0.9001–0.9466)
	Validation cohort	0.8333	0.7692	0.8824	0.9276

assessment of tumor grade in BCa tumors, and the approach showed good performance in both training and validation cohorts. In our study, by comparing the AUC values calculated by ROC curve analyses, the performance of the Joint_Model in the validation set was superior to the performance of all of the other four models, and the same trend was observed in terms of its accuracy, sensitivity, and specificity; this result suggests that the radiomics signature constructed from the Joint_Model performed best, which presents good discrimination for high-grade versus low-grade tumors, with an AUC of 0.9233 in the training cohort and 0.9276 in the validation cohort (Table 4). Although the performance of the Max-out_Model was just inferior to that of the Joint_Model, there were fewer new features introduced in the Max-out_Model. One explanation could be that to obtain a comparative estimated performance with the Joint_Model, we selected all the prominent features, which comprised the whole set of features from the three separate subsets. Perhaps the construction of the Max-out_Model could suggest a promising dimensionality reduction method that could be used in patients who are evaluated with more imaging modalities.

On the other hand, the ADC model performed better than both the T2WI and DWI models (AUCs of 0.8054 vs. 0.7828 and 0.7692, respectively). This result could be explained by the fact that both morphological T2WI and functional DWI are based on the tumor signal. The grading performance based on T2WI and DWI only used the signal intensity (SI) mean, and the SI was more applicable in differentiating the lesion from peritumoral tissues rather than in describing intratumoral characteristics and tumor grades; therefore, the SI is mostly used for BCa detection and staging rather than for grading [17, 23].

Although a previous study [17] suggested that textural features from DWI and ADC maps, especially GLCM features from ADC maps, can reflect the difference between low- and high-grade BCa, this is the first radiomics-based study regarding grade prediction of BCa to present encouraging feasibility for future use. One of the limitations of the previous study is

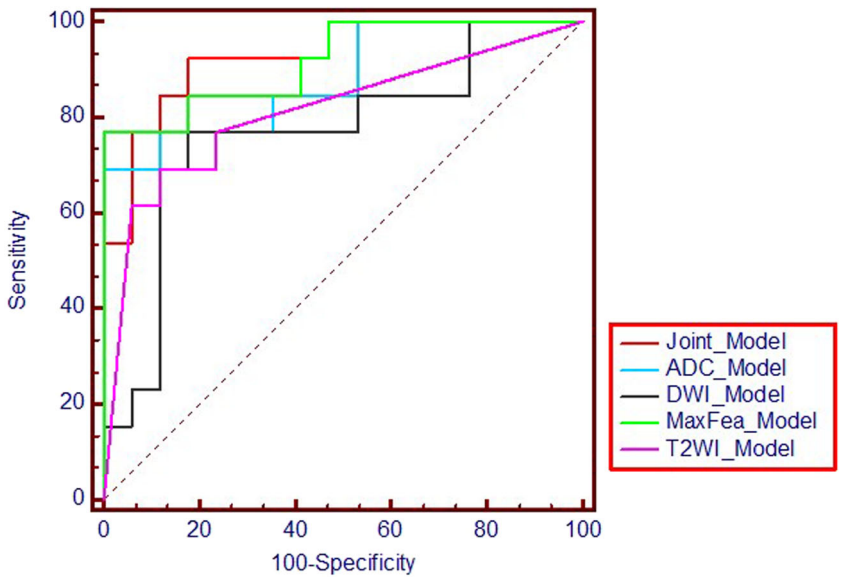

Fig. 4 ROC curves of the 5 radiomics models in the validation cohort. The ROC curves from the validation cohort show the estimated performance of the 5 models constructed in the training cohort. The AUCs of the Joint, ADC, DWI, MaxFea, and T2WI models were 0.9276, 0.8054, 0.7692, 0.9186, and 0.7828, respectively

that a validation cohort was not evaluated, and only DWI and ADC maps were analyzed. Similar to the previous study, our results also suggest that a radiomics signature based on multiparametric MRI can be applied to make noninvasive, preoperative predictions of the tumor grade of BCa. In addition, the discrimination effect was further improved in our study, with an accuracy, sensitivity, specificity, and AUC of 0.8543, 0.8267, 0.8783, and 0.9233, respectively, in the training cohort and 0.8333, 0.7692, 0.8824, and 0.9276, respectively, in the validation cohort compared with 82.9%, 78.4%, 87.1%, and 0.861, respectively, in a previous study.

Although urine cytology has been firmly established as a useful adjunct in both the diagnosis and follow-up of patients with BCa, it is highly grade-dependent. The sensitivity is low when low-grade BCa is involved, whereas the sensitivity is high in high-grade BCa, and the specificity is high [24]. Therefore, the current gold standard for the detection and surveillance of BCa is still a combination of urinary cytology with cystoscopy. However, at our institution, cytology examination was mostly performed for those patients with suspected or confirmed pelvic carcinoma or ureteral carcinoma; thus, cytology was not analyzed in our study because less than 10 BCa patients had cytology performed, and most BCa patients underwent cystoscopy and multiparametric MRI instead since this modality and the radiomics data analysis have already been reported to be feasible methods to predict tumor grade in BCa.

Our study has some limitations. First, although a validation set was included, the sample size in our study was still small, and our study was also conducted at a single hospital; thus, studies involving multiple institutions with larger sample sizes are needed. Second, radiomics signatures extracted from our MR images only contained T2WI, DWI, and ADC maps of patients who were referred clinically, and combination with other series, such as perfusion-weighted imaging (PWI) or diffusion tensor imaging (DTI), may provide more information and improve performance, since a previous study reported that quantitative parameters from PWI and DTI were significantly different in cancerous versus noncancerous areas within the bladder wall [25], and another recent study regarding diffusion kurtosis imaging suggested that kurtosis metrics performed well for differentiating NMIBCa from MIBCa [26].

In conclusion, we proposed a radiomics strategy for preoperative assessment of tumor grade in BCa, which may assist radiologists and urologists in discriminating low- and high-grade BCa tumors more precisely.

Acknowledgements The abstract of this study has been accepted as a scientific oral presentation at the European Congress of Radiology 2019 annual meeting.

Funding This study has received funding by National Natural Science Foundation of China (81701747) and Natural Science Foundation of Guangdong Province (2017A030313902).

Compliance with ethical standards

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Guarantor The scientific guarantors of this publication are Yanqiu Feng and Yan Guo.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board of the First Affiliated Hospital of Sun Yat-sen University.

Ethical approval Ethical approval was obtained from the Institutional Review Board of The First Affiliated Hospital of Sun Yat-sen University.

Methodology
- retrospective
- diagnostic or prognostic study
- performed at one institution

References

1. Siegel R, Naishadham D, Jemal A (2013) Cancer statistics, 2013. CA Cancer J Clin 63:11–30
2. Svatek RS, Hollenbeck BK, Holmang S et al (2014) The economics of bladder cancer: costs and considerations of caring for this disease. Eur Urol 66:253–262
3. Antoni S, Ferlay J, Soerjomataram I, Znaor A, Jemal A, Bray F (2017) Bladder cancer incidence and mortality: a global overview and recent trends. Eur Urol 71:96–108
4. Eble JN, Sauter G, Epstein JI, Sesterhenn IA (eds) (2004) World Health Organization classification of tumors: pathology and genetics of tumors of the urinary system and male genital organs. IARC Press, Lyon. Available via https://www.iarc.fr/wpcontent/uploads/2018/07/BB7.pdf. Accessed 2 Apr 2019
5. Linton KD, Rosario DJ, Thomas F et al (2013) Disease specific mortality in patients with low risk bladder cancer and the impact of cystoscopic surveillance. J Urol 189:828–833
6. Sylvester RJ, van der Meijden AP, Oosterlinck W et al (2006) Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. Eur Urol 49: 466–465 discussion 475–477
7. Gofrit ON, Pode D, Lazar A, Katz R, Shapiro A (2006) Watchful waiting policy in recurrent Ta G1 bladder tumors. Eur Urol 49:303–306 discussion 306–307
8. Mariappan P, Smith G (2005) A surveillance schedule for G1Ta bladder cancer allowing efficient use of check cystoscopy and safe discharge at 5 years based on a 25-year prospective database. J Urol 173:1108–1111
9. Babjuk M, Bohle A, Burger M et al (2017) EAU guidelines on nonmuscleinvasive urothelial carcinoma of the bladder: update 2016. Eur Urol 71:447–461
10. Bellmunt J, Orsola A, Leow JJ et al (2014) Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 25(suppl 3):iii40–iii48
11. Barchetti G, Simone G, Ceravolo I et al (2019) Multiparametric MRI of the bladder: inter-observer agreement and accuracy with

the Vesical Imaging-Reporting and Data System (VI-RADS) at a single reference center. Eur Radiol. https://doi.org/10.1007/s00330-019-06117-8
12. Wang HJ, Pui MH, Guo Y et al (2015) Multiparametric 3-T MRI for differentiating low-versus high-grade and category T1 versus T2 bladder urothelial carcinoma. AJR Am J Roentgenol 204:330–334
13. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446
14. Kumar V, Gu Y, Basu S et al (2012) Radiomics: the process and the challenges. Magn Reson Imaging 30:1234–1248
15. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577
16. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
17. Zhang X, Xu X, Tian Q et al (2017) Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging. J Magn Reson Imaging 46:1281–1288
18. Becker AS, Wagner MW, Wurnig MC, Boss A (2017) Diffusion-weighted imaging of the abdomen: impact of b-values on texture analysis features. NMR Biomed 30(1). https://doi.org/10.1002/nbm.3669
19. an Griethuysen JJM, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. Cancer Res 77:e104–e107
20. Vasquez MM, Hu C, Roe DJ, Chen Z, Halonen M, Guerra S (2016) Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. BMC Med Res Methodol 16:154
21. Franklin J (2010) The elements of statistical learning: data mining, inference and prediction. Publ Am Stat Assoc 99(466):567–567
22. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
23. Takeuchi M, Sasaki S, Ito M et al (2009) Urinary bladder cancer: diffusion-weighted MR imaging—accuracy for diagnosing t stage and estimating histologic grade. Radiology 251:112–121
24. Gaston KE, Pruthi RS (2004) Value of urinary cytology in the diagnosis and management of urinary tract malignancies. Urology 63:1009–1016
25. Panebianco V, De Berardinis E, Barchetti G et al (2017) An evaluation of morphological and functional multi-parametric MRI sequences in classifying non-muscle and muscle invasive bladder cancer. Eur Radiol 27:3759–3766
26. Wang F, Chen HG, Zhang RY et al (2019) Diffusion kurtosis imaging to assess correlations with clinicopathologic factors for bladder cancer: a comparison between the multi-b value method and the tensor method. Eur Radiol. https://doi.org/10.1007/s00330-018-5977-y