European Radiology (2020) 30:1804–1812
https://doi.org/10.1007/s00330-019-06484-2

ONCOLOGY

Combining DWI radiomics features with transurethral resection promotes the differentiation between muscle-invasive bladder cancer and non-muscle-invasive bladder cancer

Shuaishuai Xu[1] · Qiuying Yao[1] · Guiqin Liu[1] · Di Jin[2] · Haige Chen[2] · Jianrong Xu[1] · Zhicheng Li[3,4] · Guangyu Wu[1]

Received: 12 July 2019 / Revised: 19 September 2019 / Accepted: 9 October 2019 / Published online: 26 November 2019

Abstract

Purpose To investigate the value of radiomics features from diffusion-weighted imaging (DWI) in differentiating muscle-invasive bladder cancer (MIBC) from non-muscle-invasive bladder cancer (NMIBC).

Methods This retrospective study included 218 pathologically confirmed bladder cancer patients (training set: 131 patients, 86 MIBC; validation set: 87 patients, 55 MIBC) who underwent DWI before biopsy through transurethral resection (TUR) between July 2014 and December 2018. Radiomics models based on DWI for discriminating state of muscle-invasive were built using random forest (RF) and all-relevant (AR) methods on the training set and were tested on validation set. Combination models based on TUR data were also built. Discrimination performances were evaluated with the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, sensitivity, specificity, and F1 and F2 scores. Qualitative MRI evaluation based on morphology was performed for comparison.

Results No significant difference was found between RF and AR models. RF model was more sensitive than TUR (0.873 vs 0.655, $p = 0.019$) for discriminating muscle-invasive bladder cancer. When combining RF with TUR, the sensitivity increased to 0.964, significantly higher than TUR (0.655, $p < 0.001$), MRI evaluation (0.764, $p = 0.006$), and the combination of TUR and MRI (0.836, $p = 0.046$). Combining RF and TUR achieved the highest accuracy of 0.897 and F2 score of 0.946.

Conclusion Combining DWI radiomics features with TUR could improve the sensitivity and accuracy in discriminating the presence of muscle invasion in bladder cancer for clinical practice. Multicenter, prospective studies are needed to confirm our results.

Key Points

- *Twenty-seven to 51% of superficial bladder cancers diagnosed by transurethral resection are upstaged to muscle-invasive at radical cystectomy, suggesting its poor sensitivity for discriminating muscle-invasive bladder cancer.*
- *A small subset of selected all-relevant radiomics features exhibited an equivalent performance compared to that of all the extracted features, confirming that radiomics data contained redundant or irrelevant features and that feature selection should be performed in building radiomics models.*
- *Combining DWI radiomics features with transurethral resection could improve in clinical practice the sensitivity and accuracy for the detection of muscle invasion in bladder cancer.*

Keywords Urinary bladder cancer · Magnetic resonance imaging · Radiomics

✉ Zhicheng Li
zc.li@siat.ac.cn

✉ Guangyu Wu
danielrau@163.com

[1] Department of Radiology, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, No.1630, Dongfang road, Pudong, Shanghai 200127, China

[2] Department of Urology, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

[3] Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[4] Shenzhen Peng Cheng Laboratory, Shenzhen, China

Abbreviations

ACC	Accuracy
AR	All-relevant
AUC	Area under the receiver operating characteristic curve
BC	Bladder cancer
CIS	Carcinoma in situ
DKI	Diffusion kurtosis imaging
DTI	Diffusion tensor imaging
DWI	Diffusion-weighted imaging
GLCM	Gray-level co-occurrence matrix
GLRLM	Gray-level run length matrix
GLSZM	Gray-level size zone matrix
ICC	Intraclass correlation coefficient
MDGini	Mean Decrease in Gini index
MIBC	Muscle-invasive bladder cancer
NGTDM	Neighborhood gray-tone difference matrix
NMIBC	Non-muscle-invasive bladder cancer
PPV	Positive predictive value
RC	Radical cystectomy
RF	Random forest
ROC	Receiver operating characteristic
SEN	Sensitivity
SPE	Specificity
TUR	Transurethral resection
VOI	Volume of interest

Introduction

Preoperative differentiation between muscle-invasive bladder cancer (MIBC) and non-muscle-invasive bladder cancer (NMIBC) is crucial for subsequent treatment options. Transurethral resection (TUR) is usually chosen as the initial treatment for superficial tumors, whereas muscle-invasive tumors are treated with radical cystectomy (RC) or with adjuvant chemotherapy [1, 2]. However, 27–51% of NMIBC diagnosed by TUR are upstaged to MIBC at RC [1, 3–5], indicating its relatively low sensitivity for discriminating muscle-invasive tumors. Despite the advances in endoscopic [2] and the availability of sophisticated predicting tools [3, 6–8], accurate assessment of the clinical stage of bladder cancer (BC) is still challenging.

Magnetic resonance imaging (MRI) allows for differentiation of the bladder wall layers [9, 10]. Multiparametric MRI, including diffusion-weighted imaging (DWI), has shown promise for assessing depth of invasion in BC [11–13]. Radiomics converts medical images into mineable high-dimensional data by means of feature engineering and machine learning techniques [14–15]. Radiomics has been used to facilitate clinical decision-making in glioblastoma, lung cancer, and other solid tumors [16–18], and has shown its ability for preoperative prediction of tumor grading and lymph node metastasis in BC [19–21]. Recently, radiomics signature derived from T2WI and DWI showed potential for the differentiation of muscle invasion in BC [22, 23]. However, the sample size was relatively small and the result of TUR was not included or compared with the radiomics approach.

Thus, with a larger sample set and the result of TUR, this study aimed to develop and validate a more sensitive radiomics model from DWI for discriminating muscle-invasive bladder cancer.

Materials and methods

This study had institutional review board approval, and informed consent was waived due to its retrospective nature.

Patient population

Consecutive BC patients treated between July 2014 and December 2018 were included, according to the following criteria: (1) underwent both TUR and RC at our institute and were confirmed to have high-grade urothelial carcinoma, as almost all muscle-invasive tumors are high grade [2]; (2) delay between TUR and RC was less than 12 weeks, and absence of neoadjuvant chemotherapy or radiotherapy before RC; (3) available MRI before biopsy through cystoscopy or TUR, meaning MRI for an intact tumor. Patients were randomly divided into training set and validation set.

TUR followed by pathology investigation of the obtained specimen was a diagnostic procedure and initial treatment step. For small papillary tumors (< 1 cm), resection was performed in one piece including the part from the underlying bladder wall. For tumors > 1 cm in diameter, resection was performed in fractions including the exophytic part of the tumor, the underlying bladder wall with the detrusor muscle, and the edges of the resection area. Cauterisation was avoided as much as possible during TUR to avoid tissue deterioration. The specimen obtained by TUR was investigated by close cooperation between urologists and pathologists. The pathology report should specify tumor grade, depth of tumor invasion, presence of carcinoma in situ (CIS) or histological variant, and whether the detrusor muscle is present in the specimen. Papillary tumors confined to mucosa (Ta) or invading the lamina propria (submucosa) (T1) were classified as NMIBC. MIBC was confirmed when tumor invaded the detrusor muscle, including irregular nests, single cell infiltration, or tentacular finger-like projections.

At our institute, indications for RC included clinical MIBC and highest-risk NMIBC. Clinical highest-risk NMIBC was defined as T1HG (high grade) with any one of the following conditions or TaHG with any two: multifocal, large (> 3 cm), recurrent, associated with concurrent CIS, mixed histological variant, and BCG failure.

MR imaging

MRI including DWI for bladder was performed using a 3.0-T MR scanner (Ingenia; Philips Healthcare) with a Torso 32-channel phased array coil and without breath-holding. Parameters of DWI with single-shot EPI (echo-planar imaging) sequence were as follows: FOV, 260 × 284 × 105 mm; matrix, 132 × 170 × 32 slices; slice thickness/gap, 3/0.3 mm; TR/TE, 8216/67 ms; flip angle, 90°; number of excitations, 2; EPI factor, 71; bandwidth, 16.6 Hz; two *b* values (*b* = 0, and 1000 s/mm^2); directions of motion-probing gradients, 2; fat suppression, spectral attenuated inversion recovery; and total scan duration, approximately 2 min 50 s. Corresponding ADC maps were then automatically calculated voxel by voxel by solving the following equation:

$$S(\mathrm{b1000})/S(\mathrm{b0}) = \exp(-\mathrm{b1000} \times \mathrm{ADC})$$

where *S*(b1000) and *S*(b0) represent the signal intensity of a certain voxel in the presence and absence of diffusion sensitization, respectively.

Qualitative MRI evaluation

Invasion of muscular layer was evaluated on DWI together with T2-weighted images independently by two radiologists, according to the criteria described in [9]. Briefly, a high signal intensity tumor with a low signal intensity submucosal stalk or a thickened submucosa on DWI (*b* = 1000 s/mm^2), or an intact low signal intensity muscle layer on T2-weighted images indicated the absence of muscle invasion. For patients with multiple tumors, the one with the highest stage was documented.

Tumor segmentation

One radiologist manually segmented the entire tumor area on DWI (*b* = 1000 s/mm^2) using an open-source software package (ITK-SNAP, version 3.4.0; http://itk-snap.org) to yield volume of interest (VOI). The VOI was copied to corresponding ADC map for computer-based analysis. After 3 days, the segmentation was repeated on 40 patients by the same radiologist and by another radiologist for assessing intra- and inter-observer repeatability.

Feature extraction

The first-order intensity features, high-order texture features, and shape features were extracted within the VOIs using an in-house Matlab program (R2016a, Mathworks Inc.). The high-order texture features were extracted using several different methods, including the gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM) and neighborhood gray-tone difference matrix (NGTDM) methods. Finally, for each tumor, 156 quantitative features were extracted. Each feature was normalized into its *Z*-score.

Feature selection, and radiomics model development

Feature selection was assumed to serve as a dimension-reduction tool and discover features that may provide deeper insight to the classification task. First, intra- and inter-observer repeatability for each imaging feature was measured by intraclass correlation coefficient (ICC). Features with ICC of more than 0.85 were selected to build a classification model using random forest (RandomForest model, RF) for discriminating muscle-invasive bladder cancers. The tree number of the random forest classifier was set to 400. Mean Decrease in Gini index (MDGini) was used as variable importance measure.

For comparison, we used a random-forest based wrapper algorithm, Boruta, to select all-relevant imaging features. It evaluates feature relevance by comparing the importance of original features with that achieved by artificially added random features. Random forest is performed iteratively to measure feature importance, while irrelevant features are discarded progressively. To reach statistical significance, the algorithm repeatedly calculates all possible feature combinations, generating an all-relevant subset of features. Based on the selected all-relevant features, another random forest model (all-relevant model, AR) was built.

Combination model development

Three combination models were built. First, the result of TUR was combined with RF model and AR model, respectively, yielding two combined models. When muscle invasion was confirmed at TUR, the case was recognized as muscle-invasive regardless of the result of radiomics model. Meanwhile, if the bladder cancer was identified as non-muscle-invasive at TUR, the final result was determined based on radiomics model. For comparison, another model combining the results of TUR and qualitative MRI evaluation was also built according to the rules mentioned above.

Statistical analysis

All statistical analyses were performed using R-3.4.4 (https://www.r-project.org). All predictive models were trained on the training data set and tested on the independent validation data set. Discrimination performances were evaluated with area under the receiver operating characteristic (ROC) curve (AUC), accuracy (ACC), sensitivity (SEN), specificity (SPE), and F1 and F2 scores (F1 score is the harmonic average of the precision and recall, F2 score weighs recall higher than precision). In all tests, muscle invasion was regarded as the positive

result. Delong's test was used for comparing AUC, and McNemar's test for comparing ACC, SEN, and SPE between the two models. Inter-observer repeatability for qualitative MRI evaluation was measured by Kappa value. The R packages RandomForest and Boruta were used for model building and feature selection. All p values were two-sided. A p value < 0.05 was considered significant.

Results

Patient population

Two hundred and forty-five patients were included. After excluding 37 patients in whom radiomics features could not be extracted due to the small volume of lesions or the limited visibility of images, 218 (169 males; mean age, 66.1 years [range, 37–93]; 141 muscle-invasive tumors) were left for further analyses. In this patient group, TUR only confirmed 87 muscle-invasive tumors, and 38.3% (54/141) of RC-confirmed muscle-invasive tumors were misdiagnosed as non-muscle-invasive tumors at TUR (Table 1, Fig. 1).

Patients were randomly divided into training set (131 patients; 104 males; mean age, 65.8 years [range, 38–86]; 86 muscle-invasive tumors) and validation set (87 patients; 65 males; mean age, 66.5 years [range, 37–93]; 55 muscle-invasive tumors) (Fig. 2). No significant difference was observed in age (p = 0.696, Wilcoxon rank sum test), gender (p = 0.519, chi-square test), or muscle invasion (p = 0.824, chi-square test) between the two sets (Table 1).

Table 1 Baseline characteristics of the patients

		Total (%)	Tra set (%)	Val set (%)	p value
Number		218	131	87	
Age, year	Mean	66.1	65.8	66.5	0.696
	Range	37–93	38–86	37–93	
Gender	Male	169 (77.5)	104 (79.4)	65 (74.7)	0.519
	Female	49 (22.5)	27 (20.6)	22 (25.3)	
C stage	Non	131	80	51	0.826
	Mus	87	51	36	
P stage	Non	77	45	32	0.824
	Mus	141	86	55	
MRI	Non	101	62	39	0.823
	Mus	117	69	48	

Tra set: training set; *Val set*: validation set; *C stage*: clinical stage; *Non*: non-muscle-invasive bladder cancer; *Mus*: muscle-invasive bladder cancer; *P stage*: pathological stage

Fig. 1 Bladder MRI demonstrate a mass on the right wall of the bladder in a 79-year-old man with painless hematuria. The detrusor muscle layer seems to be intact on T2WI (**a**), and a low signal intensity thickened submucosa is observed on DWI (**b**, b value = 1000 s/mm^2), indicating the absence of muscle invasion. High-grade urothelial carcinoma staged T1 associated with concurrent carcinoma in situ is diagnosed at transurethral resection, and stratified as highest-risk non-muscle-invasive bladder cancer. Subsequently performed radical cystectomy confirmed the presence of muscle invasion

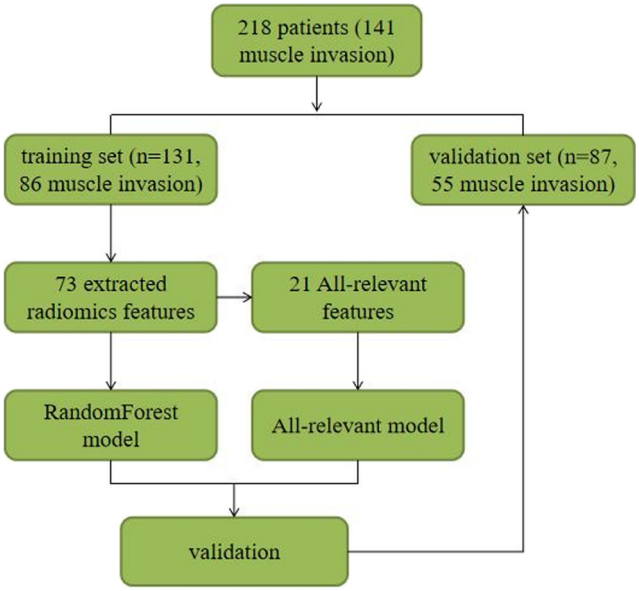

Fig. 2 Study flowchart

Radiomics and combination model development

Seventy-three features with ICC of more than 0.85 were extracted by different methods, including first order, shape, GLCM, GLRLM, GLSZM, and NGTDM features. After Boruta selection, 21 all-relevant features were obtained (Table 2) (Figs. 3 and 4). Internal validation showed no significant difference in AUC (0.907 vs 0.904, p = 0.673, Delong's test), ACC (0.839 vs 0.816, p = 0.480, McNemar's test), SEN (0.873 vs 0.855, p = 1.000), or SPE (0.781 vs 0.750, p = 1.000) between RandomForest model and all-relevant model for discriminating muscle-invasive BC (Table 3) (Fig. 5).

RandomForest model was more sensitive than TUR (0.873 vs 0.655, p = 0.019, McNemar's test), and MRI (0.873 vs 0.764, p = 0.181) for discriminating MIBC, but the difference did not reach statistical significance. When combining the RandomForest model with TUR, the sensitivity increased to 0.964, significantly higher than TUR (0.655, $p < 0.001$), MRI (0.764, p = 0.006), and the combination of TUR and MRI (0.836, p = 0.046). Notably, the combination model (RandomForest model and TUR) had the highest accuracy of 0.897 and F2 score of 0.946 for discriminating MIBC (Table 3).

Discussion

In this study, 38.3% (54/141) of RC-confirmed muscle-invasive tumors were misdiagnosed as non-muscle-invasive tumors at TUR, which is consistent with previous reports [1, 3–5]. Many reasons account for the poor sensitivity of TUR for discriminating muscle-invasive tumors, such as sampling

Table 2 A summary of 73 radiomics features with ICC of more than 0.85, and 21 all-relevant features (bold and italic) selected using Boruta

First-order features	Shape features	GLCM features		GLRLM features	GLSZM features	NGTDM features
nonText.Volume	Global1.Maximumvalue1	Uniform.GLCM.autoc	***Uniform.GLCM.inf1h***	***Uniform.GLRLM.SRE***	Uniform.GLSZM.SZE	Uniform.NGTDM.Coarseness
nonText.SurfaceArea	***Global1.Medianvalue1***	Uniform.GLCM.cprom	***Uniform.GLCM.inf2h***	Uniform.GLRLM.LRE	***Uniform.GLSZM.LZE***	Uniform.NGTDM.Contrast
nonText.Size	Global1.Minimum1	Uniform.GLCM.cshad	***Uniform.GLCM.idmnc***	Uniform.GLRLM.GLN	Uniform.GLSZM.GLN	Uniform.NGTDM.Busyness
nonText.Compactness1	Global1.Meanvalue1	***Uniform.GLCM.contr***	***Uniform.GLCM.indnc***	Uniform.GLRLM.RLN	Uniform.GLSZM.ZSN	***Uniform.NGTDM.Complexity***
nonText.Compactness2	Global1.Energy1	***Uniform.GLCM.corrm***	Uniform.GLCM.maxpr	Uniform.GLRLM.RP	***Uniform.GLSZM.ZP***	Uniform.NGTDM.Strength
nonText.Spherical_disproportion	Global1.Entropy1	***Uniform.GLCM.denth***	Uniform.GLCM.savgh	Uniform.GLRLM.LGRE	Uniform.GLSZM.LGZE	
nonText.Sphericity	Global1.Variance1	***Uniform.GLCM.dissi***	Uniform.GLCM.senth	Uniform.GLRLM.HGRE	Uniform.GLSZM.HGZE	
nonText.Ratio_surfacetovolume	Global1.Kurtosis1	Uniform.GLCM.energ	Uniform.GLCM.svarh	Uniform.GLRLM.SRLGE	Uniform.GLSZM.SZLGE	
	Global1.Root_mean_square1	Uniform.GLCM.entro	Uniform.GLCM.sosvh	***Uniform.GLRLM.SRHGE***	Uniform.GLSZM.SZHGE	
	Global1.Skewness1	***Uniform.GLCM.homom***	***Uniform.GLCM.dvarh***	Uniform.GLRLM.LRLGE	***Uniform.GLSZM.LZLGE***	
	Global1.Standard_deviation1	***Uniform.GLCM.homop***		Uniform.GLRLM.LRHGE	***Uniform.GLSZM.LZHGE***	
	Global1.Mean_absolute_deviation1			***Uniform.GLRLM.GLV***	Uniform.GLSZM.GLV	
	Global1.Uniformity1			Uniform.GLRLM.RLV	Uniform.GLSZM.ZSV	

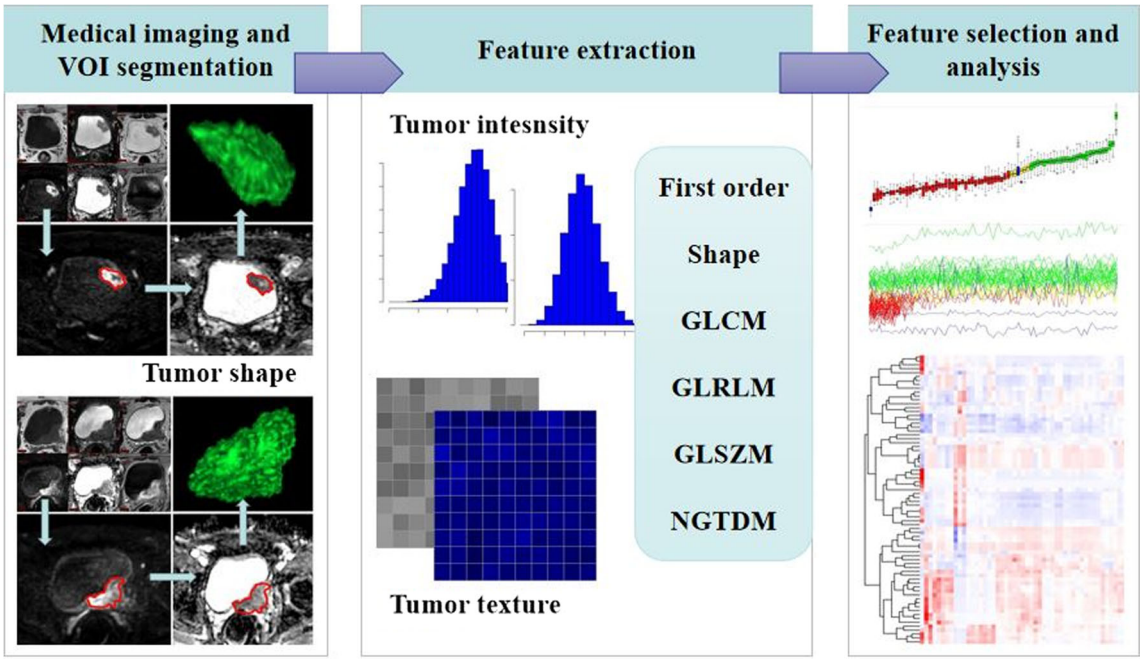

Fig. 3 Radiomics workflow

error due to incompleteness of TUR, delay in the interval from TUR to RC, and poor sensitivity of preoperative staging tools [1, 3]. Besides, qualitative MRI evaluation only showed a good inter-observer repeatability (Kappa value = 0.605) and a poor sensitivity comparable to that of TUR (0.764 vs 0.873, p = 0.181), although substantial advances in DWI have been reported to make multiparametric MRI a feasible and reasonably accurate technique to optimize the treatment of BC [9, 24].

The discrepancy between previous studies [9] and ours may be explained by the following reasons: (1) in previous report, the sample size was relatively small and the distribution of superficial and muscle-invasive tumors was uneven, leading to potential miscalculation of ACC, an imperfect evaluation index for classification performance; (2) as the authors mentioned, in cases that had underwent management before MRI, inflammatory changes due to prior treatment or biopsy may affect the results of MRI evaluation; (3) in previous report, not all patients underwent RC, and clinical stage cannot be regarded as the reference standard in the radiologic-pathologic correlation analyses due to its poor sensitivity; (4) muscle layer is usually depicted as a thin line with low signal intensity and difficult to distinguish from surrounding fat tissue on DWI. Muscle invasion can only be definitely excluded when an obvious submucosal stalk or thickened submucosa is present; otherwise, subjective judgment may lead to substantial misdiagnosis rate and poor inter-observer repeatability.

New post-processing and functional multiparametric MRI have shown promise for assessing depth of invasion in BC [11–13]. However, it is challenging to acquire images with satisfactory spatial resolution using diffusion tensor imaging (DTI) or diffusion kurtosis imaging (DKI), and these novel imaging techniques are not routinely performed in clinical practice.

Generally, there are two types of imaging features, the semantic features and the radiomics features. Semantic features are more familiar to radiologists and are commonly used to describe lesions like signal intensity or enhancement characteristics. Radiomics features are mathematically extracted quantitative descriptors, which are generally not part of the radiologists' lexicon. These features capture microscale information embedded within images, but not visible by the naked human eye [14–15]. Our radiomics model exhibited favorable discrimination performance in internal validation, with an AUC of 0.907 on the test set. The obvious advantage of TUR is its specificity of 100%, as muscle invasion is confirmed once observed at TUR specimen without considering the pathological result at RC. But for detecting highly malignant muscle-invasive BC, what physicians most importantly need is a more sensitive staging tool with a false negative rate as low as possible altogether with a relatively high positive predictive value (PPV). Recall (sensitivity) is more important than precision (PPV). Considering that F1 score is the harmonic average of the precision and recall, and that F2 score weighs recall higher than precision by placing more emphasis on false negatives, our radiomics model and combination

Fig. 4 Heatmap for normalized feature value distribution of the extracted 73 features (above) and the 21 all-relevant features (below) between superficial and muscle-invasive bladder cancers

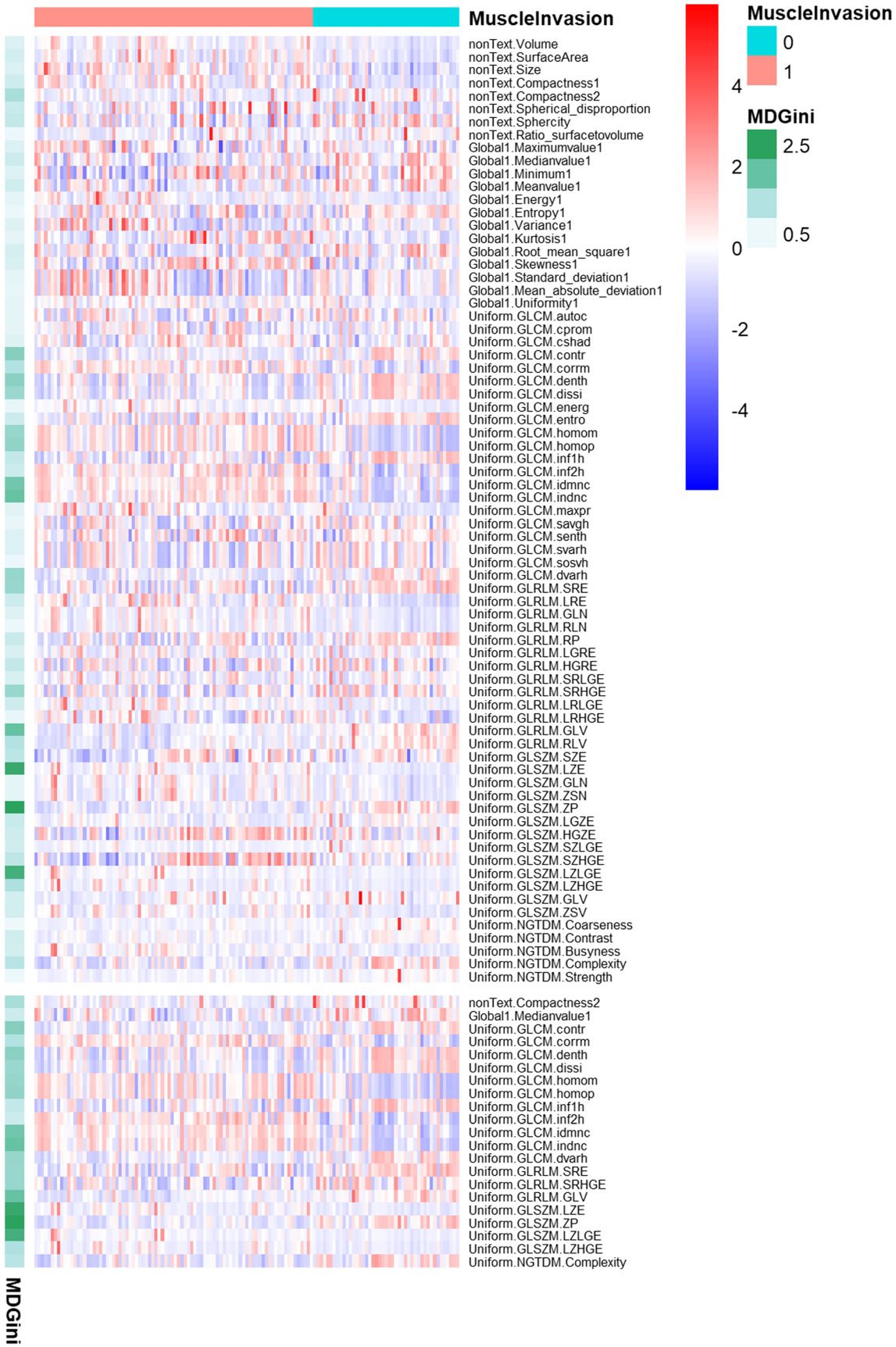
MuscleInvasion
nonText.Volume
nonText.SurfaceArea
nonText.Size
nonText.Compactness1
nonText.Compactness2
nonText.Spherical_disproportion
nonText.Sphercity
nonText.Ratio_surfacetovolume
Global1.Maximumvalue1
Global1.Medianvalue1
Global1.Minimum1
Global1.Meanvalue1
Global1.Energy1
Global1.Entropy1
Global1.Variance1
Global1.Kurtosis1
Global1.Root_mean_square1
Global1.Skewness1
Global1.Standard_deviation1
Global1.Mean_absolute_deviation1
Global1.Uniformity1
Uniform.GLCM.autoc
Uniform.GLCM.cprom
Uniform.GLCM.cshad
Uniform.GLCM.contr
Uniform.GLCM.corrm
Uniform.GLCM.denth
Uniform.GLCM.dissi
Uniform.GLCM.energ
Uniform.GLCM.entro
Uniform.GLCM.homom
Uniform.GLCM.homop
Uniform.GLCM.inf1h
Uniform.GLCM.inf2h
Uniform.GLCM.idmnc
Uniform.GLCM.indnc
Uniform.GLCM.maxpr
Uniform.GLCM.savgh
Uniform.GLCM.senth
Uniform.GLCM.svarh
Uniform.GLCM.sosvh
Uniform.GLCM.dvarh
Uniform.GLRLM.SRE
Uniform.GLRLM.LRE
Uniform.GLRLM.GLN
Uniform.GLRLM.RLN
Uniform.GLRLM.RP
Uniform.GLRLM.LGRE
Uniform.GLRLM.HGRE
Uniform.GLRLM.SRLGE
Uniform.GLRLM.SRHGE
Uniform.GLRLM.LRLGE
Uniform.GLRLM.LRHGE
Uniform.GLRLM.GLV
Uniform.GLRLM.RLV
Uniform.GLSZM.SZE
Uniform.GLSZM.LZE
Uniform.GLSZM.GLN
Uniform.GLSZM.ZSN
Uniform.GLSZM.ZP
Uniform.GLSZM.LGZE
Uniform.GLSZM.HGZE
Uniform.GLSZM.SZLGE
Uniform.GLSZM.SZHGE
Uniform.GLSZM.LZLGE
Uniform.GLSZM.LZHGE
Uniform.GLSZM.GLV
Uniform.GLSZM.ZSV
Uniform.NGTDM.Coarseness
Uniform.NGTDM.Contrast
Uniform.NGTDM.Busyness
Uniform.NGTDM.Complexity
Uniform.NGTDM.Strength
nonText.Compactness2
Global1.Medianvalue1
Uniform.GLCM.contr
Uniform.GLCM.corrm
Uniform.GLCM.denth
Uniform.GLCM.dissi
Uniform.GLCM.homom
Uniform.GLCM.homop
Uniform.GLCM.inf1h
Uniform.GLCM.inf2h
Uniform.GLCM.idmnc
Uniform.GLCM.indnc
Uniform.GLCM.dvarh
Uniform.GLRLM.SRE
Uniform.GLRLM.SRHGE
Uniform.GLRLM.GLV
Uniform.GLSZM.LZE
Uniform.GLSZM.ZP
Uniform.GLSZM.LZLGE
Uniform.GLSZM.LZHGE
Uniform.NGTDM.Complexity
MDGini
MuscleInvasion
0
1
MDGini
2.5
0.5
4
2
0
-2
-4

Table 3 A summary of the performances on validation set of RandomForest model, all-relevant model, transurethral resection, MRI, and combination models for discriminating muscle-invasive bladder cancer

	RF	AR	TUR	MRI	RF+TUR	AR+TUR	MRI+TUR
AUC	0.907	0.904	–	–	–	–	–
ACC	0.839	0.816	0.782	0.782	0.897	0.874	0.828
p value*	–	0.480	0.458	0.404	0.074	0.450	1.000
p value**	0.074	0.023	0.066	0.055	–	0.480	0.239
SEN	0.873	0.855	0.655	0.764	0.964	0.946	0.836
p value*	–	1.000	0.019	0.181	0.074	0.221	0.789
p value**	0.074	0.041	< 0.001	0.006	–	1.000	0.046
SPE	0.781	0.750	1.000	0.813	0.781	0.750	0.813
p value*	–	1.000	0.023	1.000	NA	1.000	1.000
p value**	NA	1.000	0.023	1.000	–	1.000	1.000
F1 score	0.873	0.855	0.791	0.816	0.922	0.904	0.860
F2 score	0.873	0.855	0.703	0.784	0.946	0.929	0.846

RF: RandomForest model; *AR*: all-relevant model; *TUR*: transurethral resection; *MRI*: qualitative MRI evaluation on DWI and T2WI; *AUC*: area under curve; *ACC*: accuracy; *SEN*: sensitivity; *SPE*: specificity. *p* value*: discrimination performances of different methods compared with RandomForest model; *p* value**: discrimination performances of different methods compared with the combination of RandomForest model and transurethral resection

model showed improved performance for discriminating muscle-invasive BC compared with TUR and qualitative MRI evaluation as seen on Table 3.

Another major finding of this study was that a small subset of all-relevant radiomics features selected by Boruta exhibited an equivalent performance compared to that of all the extracted features, although the classification performance using the selected optimal feature subset outperformed that using the candidate feature set in a previous report [19]. Feature selection is an important and necessary step, as it makes the model simpler and easier to interpret. When acquiring enormous amount of data ("high-dimensional"), there is an exponentially increasing risk of sparsity and loss of efficacy of traditional clustering algorithms. Feature selection addresses this issue and enhances generalization by reducing overfitting. The central premise when using a feature selection technique is that the data contains some features that are either redundant or irrelevant, and can thus be removed without incurring much loss of information [25]. Our finding suggested that radiomics data contained redundant or irrelevant features and that feature selection should be performed in building radiomics models.

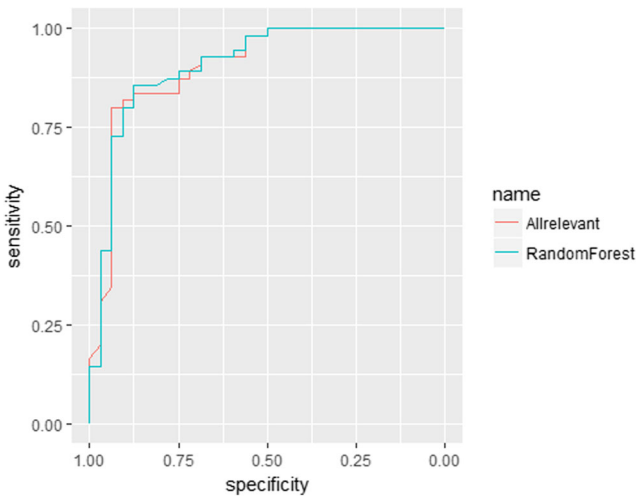

Fig. 5 ROC curves of RandomForest and All-relavant models for discriminating muscle-invasive bladder cancer on the validation set

Our study had several limitations. For cases with multiple tumors, we only documented the one with the highest stage for radiologic-pathologic correlation analyses. Although each tumor was respectively analyzed in previous report [9], our method was closer to clinical practice. Incorrect manual segmentation, because either of the small volume of the lesions or of the limited visibility of the images, may lead to poor repeatability of feature extraction. So some ineligible cases were excluded. Moreover, external validation for the radiomics model was not performed. In the future, multicenter validation with a larger sample size is needed to acquire high-level evidences.

In conclusion, a radiomics model from DWI was more sensitive and accurate than TUR and could help for discriminating muscle-invasive bladder cancer in clinical practice. Multicenter, prospective studies are needed to confirm our results.

Acknowledgements The authors thank their colleagues of the department of radiology of their institute.

Funding information This study has received funding by the National Natural Science Foundation of China; contract grant numbers are the following: Youth Program Nos. 81601487 and 81672514.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Guangyu Wu.

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.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- Retrospective
- Diagnostic or prognostic study
- Performed at one institution

References

1. Babjuk M, Böhle A, Burger M et al (2017) EAU guidelines on non-muscle invasive urothelial carcinoma of the bladder: update 2016. Eur Urol 71:447–461
2. Alfred Witjes J, Lebret T, Compérat EM et al (2017) Updated 2016 EAU guidelines on muscle-invasive and metastatic bladder cancer. Eur Urol 71:462–475
3. Karakiewicz PI, Shariat SF, Palapattu GS et al (2006) Precystectomy nomogram for prediction of advanced bladder cancer stage. Eur Urol 50:1254–1260
4. Shariat SF, Palapattu GS, Karakiewicz PI et al (2007) Discrepancy between clinical and pathologic stage: impact on prognosis after radical cystectomy. Eur Urol 51:137–149 discussion 49-51
5. Svatek RS, Shariat SF, Novara G et al (2011) Discrepancy between clinical and pathological stage: external validation of the impact on prognosis in an international radical cystectomy cohort. BJU Int 107:898–904
6. Shariat SF, Margulis V, Lotan Y, Montorsi F, Karakiewicz PI (2008) Nomograms for bladder cancer. Eur Urol 54:41–53
7. Green DA, Rink M, Hansen J et al (2013) Accurate preoperative prediction of non-organ-confined bladder urothelial carcinoma at cystectomy. BJU Int 111:404–411
8. Shariat SF, Passoni N, Bagrodia A et al (2014) Prospective evaluation of a preoperative iomarker panel for prediction of upstaging at radical cystectomy. BJU Int 113:70–76
9. 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
10. Green DA, Durand M, Gumpeni N et al (2012) Role of magnetic resonance imaging in bladder cancer: current status and emerging techniques. BJU Int 110:1463–1470
11. Lee M, Shin SJ, Oh YT et al (2017) Non-contrast magnetic resonance imaging for bladder cancer: fused high b value diffusion-weighted imaging and T2-weighted imaging helps evaluate depth of invasion. Eur Radiol 27:3752–3758
12. 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
13. 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 29:4447–4455
14. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
15. Gillies RJ, Kinahan PE, Hricak H et al (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577
16. Kotrotsou A, Zinn PO, Colen RR (2016) Radiomics in brain tumors: an emerging technique for characterization of tumor environment. Magn Reson Imaging Clin N Am 24:719–729
17. Lee G, Lee HY, Park H et al (2017) Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: state of the art. Eur J Radiol 86:297–307
18. Li ZC, Zhai G, Zhang J et al (2019) Differentiation of clear cell and non-clear cell renal cell carcinomas by all-relevant radiomics features from multiphase CT: a VHL mutation perspective. Eur Radiol 29:3996–4007
19. 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 6:1281–1288
20. Wang H, Hu D, Yao H et al (2019) Radiomics analysis of multiparametric MRI for the preoperative evaluation of pathological grade in bladder cancer tumors. Eur Radiol. https://doi.org/10.1007/s00330-019-06222-8
21. Wu S, Zheng J, Li Y et al (2017) A radiomics nomogram for the preoperative prediction of lymph node metastasis in bladder cancer. Clin Cancer Res 23:6904–6911
22. Xu X, Liu Y, Zhang X et al (2017) Preoperative prediction of muscular invasiveness of bladder cancer with radiomic features on conventional MRI and its high-order derivative maps. Abdom Radiol (NY) 42:1896–1905
23. Xu X, Zhang X, Tian Q et al (2019) Quantitative identification of nonmuscle-invasive and muscle-invasive bladder carcinomas: a multi parametric MRI radiomics analysis. J Magn Reson Imaging 49:1489–1498
24. Verma S, Rajesh A, Prasad SR et al (2012) Urinary bladder cancer: role of MR imaging. Radiographics 32:371–387
25. Bermingham ML, Pong-Wong R, Spiliopoulou A et al (2015) Application of high-dimensional feature selection: evaluation for genomic prediction in man. Sci Rep 5:10312