Computers in Biology and Medicine 144 (2022) 105318

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

The potential of prostate gland radiomic features in identifying the Gleason score

Lixin Gong [a,b], Min Xu [c], Mengjie Fang [b,d], Bingxi He [b,e], Hailin Li [b,e], Xiangming Fang [c,***], Di Dong [b,d,**], Jie Tian [a,b,e,f,*]

[a] College of Medicine and Biological Information Engineering School, Northeastern University, Shenyang, 110016, China
[b] CAS Key Laboratory of Molecular Imaging, The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China
[c] Imaging Center, Wuxi People's Hospital, Nanjing Medical University, Wuxi, 214023, China
[d] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, 100049, China
[e] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, 100191, China
[f] Zhuhai Precision Medical Center, Zhuhai People's Hospital (affiliated with Jinan University), Zhuhai, 519000, China

ARTICLE INFO

Keywords:
MRI
Radiomics
Prostate cancer
Gleason score
Gland segmentation

ABSTRACT

Background: Gleason score (GS) is one of the most critical predictors of diagnosing prostate cancer (PCa). The prostate gland, including both lesions and their microenvironment, may contain more comprehensive information about the PCa. We aimed to investigate the potential of prostate gland radiomic features in identifying Gleason scores (GS) < 7, $= 7$, and >7. Methods: We retrospectively examined preoperative magnetic resonance imaging (MRI) results, clinical data, and postoperative pathological findings from 489 PCa patients. The three-dimensional (3D) and two-dimensional (2D) radiomic features were extracted from the manually segmented 3D prostate gland and its maximum 2D layer on MRI, respectively. Significant features were selected, and sequence signatures were then developed via multi-class linear regression (MLR) accordingly. Subsequently, 2D and 3D radiomic models were constructed by applying MLR to the combination of the sequence signatures, respectively. The stability of the significant features was discussed by their average ranking in the other 30 random cohorts. Based on our distance matrix algorithm, we generated different regions of interest to simulate the manual segmentation biases and discuss the model's tolerance to them. Results: Our 2D model reached a C-index of 0.728 and an average area under the receiver operating characteristic curve of 0.794 in the validation cohort. The corresponding key features were stable, with an average ranking of the top 8.352% in 30 random cohorts, and the model could tolerate a segmentation boundary deviation of 2 mm without significant performance degradation. Conclusion: 2D prostate-gland-MRI-based radiomic features showed stable potential in identifying GS.

1. Introduction

Prostate cancer (PCa) ranked second in the fatal cancer causes for men in 2019 as the most commonly diagnosed male cancer in the United States [1,2]. The previous study reported that early detection of asymptomatic PCa could reduce mortality by 20% [3,4].

National Comprehensive Cancer Network [5] demonstrated that the Gleason score (GS) is one of the most critical predictors of diagnosing PCa, according to which patients can be divided into three prognosis groups with GS $<$ 7, GS $=$ 7, and GS $>$ 7. Previous studies have

Abbreviations: PCa, prostate cancer; GS, Gleason score; MRI, magnetic resonance imaging; ROI, region of interest; DWI, diffusion-weighted images; ADC, apparent diffusion coefficient; T2WI, T2-weighted images; CM, clinical model; PGRFM, model based on prostate gland radiomic features; MLR, multi-class linear regression; ROC, the receiver operating charasteristic; AUC, the area under ROC curve; SP, specificity; SE, sensitivity.

* Corresponding author. CAS Key Laboratory of Molecular Imaging, the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.

** Corresponding author. CAS Key Laboratory of Molecular Imaging, the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.

*** Corresponding author. Imaging Center, Wuxi People's Hospital, Nanjing Medical University, China.

E-mail addresses: xiangming_fang@njmu.edu.cn (X. Fang), di.dong@ia.ac.cn (D. Dong), jie.tian@ia.ac.cn (J. Tian).

https://doi.org/10.1016/j.compbiomed.2022.105318
Received 4 November 2021; Received in revised form 8 February 2022; Accepted 9 February 2022
Available online 11 February 2022

consistently demonstrated that GS is closely related to PCa progression, outcomes, and mortality [6]. Relatively conservative treatment strategies are beneficial for PCa patients with GS < 7 [7], and months of neoadjuvant androgen deprivation therapy (ADT) may be recommended in patients with GS $= 7$ [8], while those with GS > 7 would consider receiving neoadjuvant ADT for more than two years [9]. Therefore, it is essential to accurately determine the GS before surgery. However, GS can only be determined via needle biopsy or surgical pathology. The transrectal and transperineal approaches are the two primary biopsy methods for the diagnosis of PCa [10]. The transrectal approach is more frequently performed by urologists, however, it was been associated with inadequately sampling anterior prostate regions, high false-negative rates, and high infection rates [11,12]. It has been shown that the transperineal approach can help improve sampling of the anterior and apical regions in case of repeat biopsy with lower infection rates [13,14]. Nevertheless, more than 68% of PCa was multi-lesioned and had scattered distribution in the prostate gland, especially for early PCa with a small volume showing multifocal, heterogeneous, and dispersed growth. The local biopsy cannot reflect the overall situation of the PCa. Whether increasing the number of cores equates to a higher cancer yield is debatable, but both the transrectal and transperineal biopsies approaches would tend to have increasing potential complications of pain, erectile dysfunction, and urinary retention [15]. Therefore, non-invasive and precise GS prediction methods are needed.

Research has indicated that standard magnetic resonance (MR) imaging (MRI) findings are correlated with GS obtained via pathological analysis [16–18]. Multiparametric MRI (mpMRI), including T1-weighted imaging, T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE) MRI, has emerged as an important tool in the early diagnosing of PCa [19–21] and is usually considered as the best imaging method for evaluating PCa [22,23]. Besides, considering the controversial value and high cost of DCE MRI for identifying PCa, biparametric MRI (bpMRI) protocol that only employs the T2WI and DWI sequences has also been generalized applied [24,25]. Several studies have demonstrated that radiomic analysis, which extracts large amounts of quantitative MRI information, can be used to identify PCa, distinguish GS, and determine aggressiveness [17,18,26,27]. However, it's a pity that these previous studies relied on complex lesion region segmentation, which was tended to ignore scattered and obscure lesions and be time-consuming.

A series of studies, getting rid of the complex segmentation of the tumor lesions and concentrating on the prostate gland, demonstrated the predictive values of the information about the prostate gland for GS, such as the volume of the whole prostate gland (VPG), and so on [28,29]. Compared with the tumor region, the prostate gland includes both lesions and their microenvironment, which may contain more comprehensive information about the tumor considering the characteristics of multiple lesions and scattered distribution of PCa. There have been studies analyzing the whole organ instead of a specific lesion, such as a whole lung radiomic analysis on the chest radiograph [30,31]. Inspired by these studies, we attempted to explore the ability of the prostate gland as the region of interest (ROI) in GS prediction and constructed a radiomic predictive model. Meanwhile, we introduced a distance matrix based algorithm for simulating the manual segmentation deviation and investigating the model's anti-interference ability against it.

2. Materials and methods

2.1. Patients and MRI data acquisition

We collected patients with PCa treated at Wuxi People's Hospital from June 2008 to January 2018. The requirement for informed consent was waived by the Ethics Committee of the hospital.

All patients were scanned at a single institution with 3.0-T MR scanners (MAGNETOM Verio, Siemens Healthcare, Erlangen, Germany) using pelvic phased-array coils. To guarantee the image quality, the MR machine was checked on by a hospital radiological technician every month and was further maintained by a Siemens engineer on a bimonthly basis. For every standard prostatic MRI examination conducted in Wuxi People's Hospital, three types of MRI data were evaluated: T2-weighted imaging (T2WI), Diffusion-weighted imaging (DWI) and its derivative apparent diffusion coefficient (ADC) imaging. In our study, the MRI scanning was guided to be strictly consistent with the latest guidelines [32] with the scanned b-values of series values from 0 to high value step-by-step (from 2008 to 2015: b-values were 0, 50, 100, 150, 200, 400, 600 and 800 s/mm^2; from 2015 to 2018: b-values were 0, 50, 100, 150, 200, 600, 800 and 1600 s/mm^2). The present study used a b-value of 800 s/mm^2 uniformly, considering enrolling as much as patients to help reduce the randomness of experimental results.

Inclusion criteria were as follows: (a) availability of pathological results and preoperative MRI data; (b) interval between image capture and pathological examination of no more than six weeks. (c) Gleason scores of different lesions (from systemic puncture with 12+X needle biopsy, or from postoperative pathology of patients underwent radical prostatectomy) were the same. The exclusion criteria were as follows: (a) patients whose critical clinical information was unavailable; (b) the MRI was too unclear to delineate the boundary of the prostate gland.

Under the constraints of inclusion and exclusion criteria, a total of 489 eligible patients were finally studied. In our study, 2/3 of the total samples were randomly selected as the primary cohort (n = 326), while the remaining patients were included in the validation cohort (n = 163). The study design is shown in Fig. 1.

2.2. Clinical characteristics

Since that VPG has been reported to be related to PCa, we added it as a potential clinical characteristic to be analyzed. Our candidate clinical predictors thus included age, prostate-specific antigen (PSA) levels, the density of PSA (PSAD), and VPG. Among the parameters derived from serum PSA values [33], PSAD possesses the greatest diagnostic efficacy and is closely related to the pathological severity of aggressive/high-grade PCa [34] (Supplement I). In our study, the values of PSAD were calculated according to the values of PSA and VPG.

$$v_{PSAD} = \frac{v_{PSA}}{v_{VPG}} \tag{1}$$

where v_{PSAD} and v_{PSAD} are the values of PSAD and PSA, respectively. v_{VPG} is the value of VPG obtained in T2WI [35].

We used the non-parameteric test to retain the significant characteristics to construct a clinical model (CM) for predicting GS via the multi-class linear regression (MLR).

2.3. Region of interest masking and image resampling

An experienced radiologist manually denoted the entire prostate gland as a ROI in all three types of 3D MR images utilizing ITK-SNAP version 3.4.0 (www.itksnap.org). The prostate gland was roughly outlined according to the Prostate Imaging - Reporting and Data System (PIRADS) v2 [19], and anatomical landmarks such as the urethra, ejaculatory ducts, and prostatic capsule were used to determine the ROI, while the surrounding fat, muscle tissue, large blood vessels, and veins were excluded. We also performed the interrater and intrarater reliability studies, which showed differences between different observers had little influence on our study (Supplement II. Manual Segmentation Evaluation).

Since 3D and 2D features may lead to a difference in the predictive performance, we extracted 3D radiomic features from 3D ROIs and 2D radiomic features from the 2D ROIs that is the single transverse MRI slice covering the largest area of the 3D ROI [36]. It is worth noting that the pixel spacing of the ROI was normalized with an average spacing via three order B-Spline interpolation to obtain the isotropic volumes. The

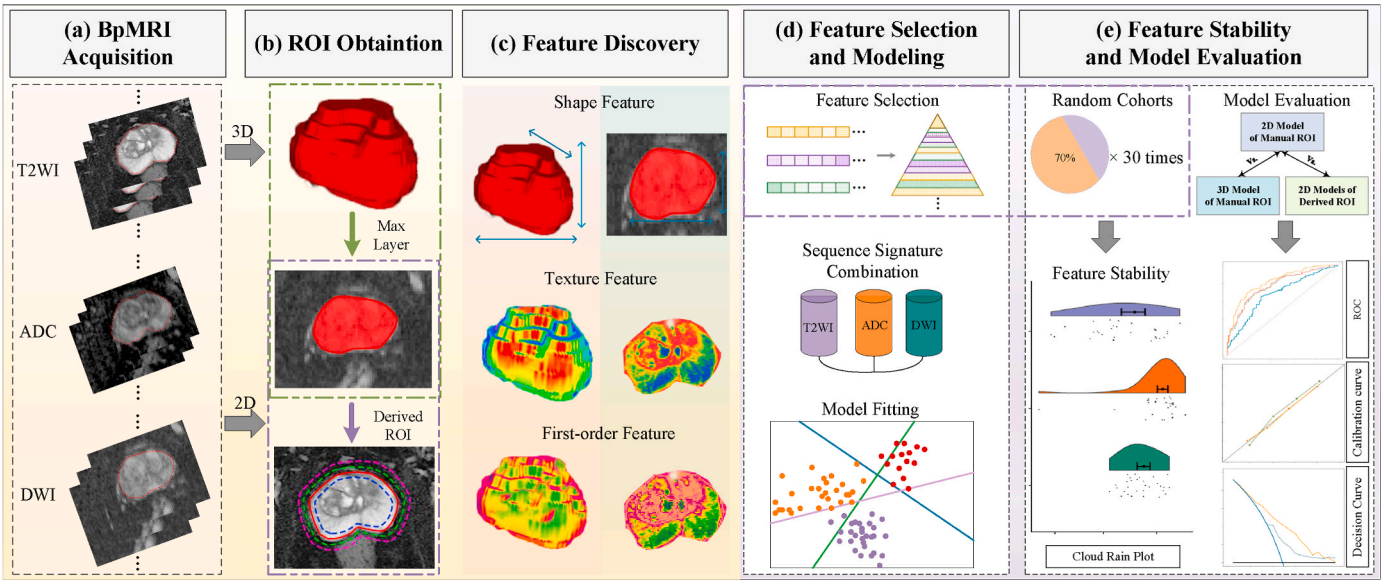

Fig. 1. Exploring the potential predictive power of the prostate gland radiomic features for GS. (a) MRI acquisition. (b) ROI obtaintion, including the 3D manual segmented ROI by radiologists, the 2D segmentation, and the generated ROIs with a distance of 1 mm and 2 mm. (c) 3D and 2D radiomic features extracted from the 3D ROI and 2D ROI, respectively, including shape, texture, and first-order features. (d) Features were selected, and the sequence-level signatures were constructed, based on which models were further developed via linear regression analysis. (e) The stability of key features was validated in the other 30 random primary cohorts and exhibited in the raincloud plot. 3D model, 2D model of the manual ROI, and the model of the generated ROI were evaluated using ROC curves, decision curves, and calibration curves. ROI: Region of the interest; PCa: prostate cancer; MRI: magnetic resonance imaging; ROC: receiver operating characteristic; 2D: two-dimensional; 3D: three-dimensional.

average pixel spacings in the x, y, and z directions were 0.629 mm/voxel, 0.629 mm/voxel, and 3.909 mm/voxel in the T2WI sequence, respectively, and 1.522 mm/voxel, 1.522 mm/voxel, and 4.539 mm/voxel in the DWI and ADC sequences, respectively. Accordingly, the 3D ROIs were resampled to the isotropic ones with the target pixel spacing of 1 mm/voxel in the T2WI sequence and 2 mm/voxel in DWI and ADC sequences, while the 2D ROIs were resampled with the target pixel spacing of 0.63 mm/voxel in the T2WI sequence and 1.52 mm/voxel in the DWI and ADC sequences.

2.4. Feature discovery and sequence signature building

Radiomic features were extracted using the package of Pyradiomics [37] in Python (version 3.6; https://www.python.org/) for Windows. For each MRI sequence of each case, we extracted three groups of features from ROIs, including the first-order, shape, and textural features, which have been used in several previous studies [38–41]. First-order features calculated from the histogram of all tumour intensities described the intensity characteristics, such as energy, entropy, and so on. Shape features quantitatively described the 3D size and shape of the tumour. Texture features were based on the quantification of tumour heterogeneity.

In this study, to make the datasets have a unified order of the magnitude, features in both the primary and validation cohorts were scaled using the Z-score technique with the same mean and standard deviation calculated from the primary cohort. Meanwhile, the null and useless values were also processed. The detailed data pre-process is shown in the section "Supplement III. Data Pre-Process".

Radiomic features showing strong correlations with GS need to be selected. Considering the wide variety of types and properties of the radiomic features, we applied the non-parameteric Kruskal-Wallis test [42], the minimum redundancy-maximum relevance (mRMR) test [43], and the sequential backward elimination (SBE) [44] algorithm performed within the MLR analysis to select the key features. Specifically, features with p-values less than 0.05 in the Kruskal-Wallis test, rankings of top 30 in the mRMR test [45], and p-values less than 0.05 in the multivariate analysis employed with the SBE algorithm were selected to construct the sequence signatures. Then, we analyzed the selected key features to develop the radiomic single sequence signature using MLR as follows [46].

$$S_k = \sum_{i=1}^{N_k} w_{k,i} F_{k,i} + v_k, \mathrm{k} = 1, 2, 3 \quad (2)$$

where S_k is the sequence signature constructed from the k_{th} sequence of MRI (k_1: T2WI, k_2: ADC, k_3: DWI), $F_{k,i}$ is the i_{th} significant radiomic feature of the k_{th} MRI sequence, and $w_{k,i}$ is the coefficient for it. N_k and v_k is the number of significant features and the bias of the k_{th} MRI sequence.

2.5. Model construction and validation

Based on the sequence signatures constructed from T2WI, ADC, and DWI, radiomic models were constructed via MLR. Specifically, a 2D radiomic model (2D-PGRFM) and a 3D radiomic model (3D-PGRFM) were correspondingly constructed employing 2D or 3D radiomic sequence signatures. The formula is as follows:

$$M = \sum_{k=1}^{3} \lambda_k S_k + \beta \quad (3)$$

where M is the radiomic model, S_k is the sequence signature constructed from the k_{th} sequence of MRI, λ_k and β are the coefficient and bias for the sequence signature S_k, which is determined via the MLR.

It is noticed that all the strategies used for 3D and 2D PGRFM are the samely designed during the feature extraction, selection and model construction, respectively. In the same way, the combined model integrating both the radiomic key features and clinc information was also constructed to explore the radiomic model's benefit for the clinical information.

C-index was calculated to evaluate the performance of the model on identifying GS < 7 vs. GS $= 7$ vs. GS > 7. To further evaluate the model's performance, logistic regression (LR) analyses were used to show the predictive performance of the radiomic models for GS subgroups (GS < 7 vs. GS $\geqslant 7$, GS $\leqslant 7$ vs. GS > 7) based on the current three-way classification results (Fig. 2). The area under the receiver operating

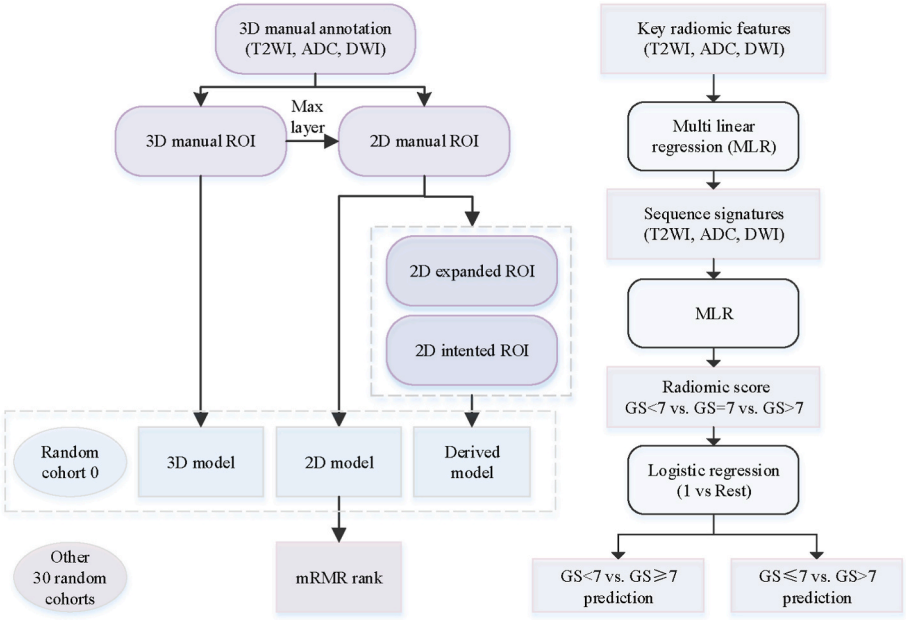

Fig. 2. The technical route of the radiomic construction and evaluation.

characteristic (ROC) curve (AUC), specificity (SP), sensitivity (SE), and accuracy (ACC) values were reported separately in the two subgroups of the GS, and the optimum cutoff point was determined in the primary cohort using the Youden index. To validate the robustness of the radiomic models, the corresponding quantitative metrics were calculated in the validation cohort. We also analyzed the model performance using decision curve analysis (DCA) and calibration curves, and the former aimed to identify the maximum net benefit while the latter reflected the performance of each model according to the consistency between the predicted and observed GS.

2.6. Feature stability in random datasets

The partitioning of the cohorts may influence unstable radiomic features. To explore the stability of our key radiomic features utilized for PGRFM, we randomly split the whole dataset into 30 pairs of the primary and validation cohorts (clinical characteristics had no significant differences between them). For these 30 random primary cohorts, we performed the same Kruskal-Wallis test and mRMR as we used for our PGRFM, after which the features rankings were acquired. Accordingly, the matching rankings of our PGRFM key radiomic features were summarized and calculated as follows.

$$R_{i,k} = \frac{1}{N_k}\sum_{j=0}^{N_k}\frac{R_{i,j,k}(x)}{M_{i,k}} \tag{4}$$

$$R_k = \frac{1}{30}\sum_{i=0}^{30}R_{i,k} \tag{5}$$

where N_k is the number of significant features in the k_{th} sequence, $R_{i,j,k}(x)$ is the ranking of the j_{th} significant feature of the k_{th} sequence signature in the i_{th} random primary cohort, $M_{i,k}$ is the number of selected features participating in mRMR ranking for the k_{th} sequence in the i_{th} random primary cohort, $R_{i,k}$ is the average ranking of the k_{th} MRI sequence key features in the i_{th} random primary cohort, and R_k is the average ranking of the k_{th} MRI sequence features within the 30 random primary cohorts.

We also used the raincloud plots [47] to exhibit the ranking distributions for each key feature and their average rankings in each sequence. The raincloud plot could show the data distribution with a cloud diagram showing nuclear density distribution, a marker showing the median of the sample, a marker indicating the interquartile range, and all the sample points.

2.7. Tolerance for segmentation deviation

Subjective differences are common and inevitable for manual segmentation, even for simpler gland segmentation than lesion segmentation. To explore the tolerance of our model for gland segmentation deviation, we introduced a distance matrix based algorithm, via which we obtained the automatically derived expanded and indented ROIs according to the required pixel distance to the original radiologist's manually segmented ROI on the maximum MRI layer. The algorithm details are as follows:

1. Obtain the distance matrix between the original manually segmented ROI and its surrounding pixels.

$$R_m = \left\{\left(x_k^R, y_k^R\right) | k = 0, 1, \ldots, \mathcal{N}_{ROI}\right\} \tag{6}$$

$$I_m = \{(x,y) | 0 \le x < \mathcal{W}, 0 \le y < \mathcal{H}\} \tag{7}$$

where R_m is the original manually annotated ROI on the maximum layer I_m, $\mathcal{N}_{ROI}$ is the number of pixels in R_m, $\left(x_k^R, y_k^R\right)$ is the coordinate of the k_{th} pixel in R_m, $\mathcal{W}$ and $\mathcal{H}$ represent the width and height of the area around the I_m to be analyzed.

Taking into the variations in pixel spacing among MRIs, we calculated euclidean norm rather than the physical distance.

$$d(x,y) = \min\left\{\left\|(x,y), \left(x_k^R, y_k^R\right)\right\|_2\right\}, \quad k = 0, 1, \ldots, \mathcal{N}_{ROI} \tag{8}$$

where $d(x, y)$ is the distance between pixel (x, y) and R_m, which was defined as the minimum euclidean norm of (x, y) to all the pixels in R_m.

2. Obtain the derived ROIs according to the required physical distance, such as 1 mm, 2 mm, etc. Taking into the variation in pixel spacing between MRIs, we first convert the required physical distance to the required number of pixels.

$$D_{pixel} = \frac{D_{physical}}{S_{avg}} \tag{9}$$

$$\Delta_{expand} = \{(x,y)|d(x,y) \leq 2D_{pixel} \wedge (x,y) \notin R_m\} \quad (10)$$

$$\Delta_{intend} = \{(x,y)|d(x,y) \leq 2D_{pixel} \wedge (x,y) \in R_m\} \quad (11)$$

$$R_{expand} = R_{\mathrm{m}} \cup \Delta_{expand} \quad (12)$$

$$R_{intend} = R_{\mathrm{m}} - \Delta_{intend} \quad (13)$$

where $D_{physical}$ is the required physical distance to be expanded or intended to R_m, S_{avg} is the average pixel spacing, D_{pixel} is the corresponding required pixel distance, R_{expand} is the expanded ROI by expanding region Δ_{expand} based on the required distance to the R_m, and R_{intend} is the indented ROI by indenting region Δ_{intend} based on the required distance to the R_m.

2.8. Influence of unbalanced data distribution

Our datasets had an unbalanced distribution between the GS < 7 and GS $\geqslant 7$ subgroups, which is common for clinic data and may have an influence on our model's predictive performance. Therefore, we conducted the random oversampling (ROS) on the primary cohort to explore the model performance benefit from the balanced datasets.

2.9. Statistical analysis

Chi-squared and Kruskal-Wallis tests were used to assess the differences among the discrete and continuously variable groups, respectively. Two-sided statistical tests were conducted, and statistical significance was determined when the p-value was less than 0.05. R software for Windows (version 4.0.5; https://www.r-project.org) was applied for statistical analyses.

3. Results

3.1. Clinical characteristics and CM performance

Patients' clinical characteristics are summarized in Table 1. As shown in Table 1, PSAD and PSA exhibited $p < 0.05$ in both the primary and validation cohorts and were used to analyze the predictive power of the CM. However, age and VPG were not used since they were identified as insignificant factors (primary cohort: age: $p = 0.355$, volume: $p = 0.155$; validation cohort: age: $p = 0.988$, volume: $p = 0.492$).

The CM, including PSA and PSAD, exhibited poor performance in both the primary and validation cohorts. Specifically, it got a C-index of 0.682, an average AUC of 0.771, and an average ACC of 0.639, an average specificity of 0.738, and an average sensitivity of 0.623 (GS < 7 vs. GS $\geqslant 7$: AUC = 0.790, specificity = 0.800, sensitivity = 0.595; GS $\leqslant 7$ vs. GS > 7: AUC = 0.752, specificity = 0.675, sensitivity = 0.651) in the validation cohort (Table 2). The ROC curves for CM of GS subtypes are shown in Fig. 3.

3.2. Feature discovery and radiomic signature construction

The initial 3D radiomic feature set contained 4,227 features from the manually segmented ROIs, including 1,409 features extracted in each MRI sequence: 14 shape features, 1,125 textural features, and 270 first-order statistical features. The initial 2D radiomic feature set contained 3,138 features from the maximum layer of the manually segmented ROIs, including 1,046 features in each MRI sequence: 23 shape features, 825 textural features, and 198 first-order statistical features.

After the feature selection, 3, 5, 6 key features were determined for T2WI, ADC, and DWI of 3D ROIs respectively, while 3, 3, 4 key features were correspondingly determined for 2D ROIs. The detailed key features are shown in Supplement Table A1. The calculation of the sequence signatures and radiomic score of models are shown in Supplement IV.

3.3. PGRFM and predictive potential

For 3D-PGRFM, the C-index reached 0.753 (0.793–0.832), an average ACC of 0.729, and an average AUC of 0.800 with an average specificity of 0.781 and an average sensitivity of 0.703 in the primary cohort, and it obtained a C-index of 0.706 (0.767–0.827), an average ACC of 0.724, and an average AUC of 0.781 with an average specificity of 0.704 and an average sensitivity of 0.727 in the validation cohort in predicting GS < 7 vs. GS $= 7$ vs. GS > 7 (Table 2, Fig. 3).

For 2D-PGRFM, the C-index reached 0.714 (0.757–0.800), an average ACC of 0.701, and an average AUC of 0.776 with an average specificity of 0.798 and an average sensitivity of 0.674 in the primary cohort, and it earned a C-index of 0.728 (0.784–0.840), an average ACC of 0.727, and an average AUC of 0.794 with an average specificity of 0.765 and an average sensitivity of 0.743 in the validation cohort in predicting GS < 7 vs. GS $= 7$ vs. GS > 7, as shown in Table 2 and Fig. 3.

When compared with CM, both 2D-PGRFM and 3D-PGRFM had statistically different improvements in identifying GS $\leqslant 7$ vs. GS > 7 ($p < 0.05$), although insignificant differences were found in identifying GS < 7 vs. GS $\geqslant 7$. The calibration curves and decision curves for PGRFM are shown in Fig. 4. The calibration curves of PGRFM predicting GS of the patients with prostate cancer demonstrated good agreements between the observations and predictions in both the primary and validation

Table 1
Clinical characteristics of patients in the primary and validation cohorts.

Characteristics	Primary cohort					Difference between cohorts (p-value)	Validation cohort				
	GS < 7	GS $= 7$	GS > 7	mean ± SD	significance analysis (p-value)		GS < 7	GS $= 7$	GS > 7	mean ± SD	significance analysis (p-value)
No. of patients	38	136	152			0.398	15	62	86		
Age (y)		–		72.479 ± 7.447	0.355	0.630		–		73.411 ± 6.643	0.988
VPG (mL)		–		35.565 ± 20.234	0.155	0.513		–		35.018 ± 17.621	0.492
PSA ($< 10ng/mL$)	16	31	15	94.653 ± 350.985	*	0.533	6	11	8	82.815 ± 276.488	*
PSA ($\geqslant 10ng/mL$)	22	105	137				9	51	78		
PSAD ($< 0.45ng/mL^2$)	23	45	28	2.951 ± 9.193	*	0.507	10	18	10	2.269 ± 4.952	*
PSAD ($\geqslant 0.45ng/mL^2$)	15	91	124				5	44	76		

GS: Gleason score; SD: standard deviation; PSA: prostate-specific antigen; PSAD: PSA density; VPG: volume of the prostate gland; *: p-value $<$ 0.05; Chi-squared and Kruskal-Wallis tests were used to assess the differences among the discrete and continuous variable groups, respectively.

Table 2
Model performance on the manual ROI.

Evaluation	GS < 7 vs. GS = 7 vs. GS > 7	GS < 7 vs. GS ⩾ 7				GS ⩽ 7 vs. GS > 7			
	C-index	AUC	SP	SE	ACC	AUC	SP	SE	ACC
CM									
Primary cohort	0.625 (0.677–0.729)	0.712 (0.620–0.804)	0.790	0.552	0.580	0.683 (0.625–0.741)	0.695	0.651	0.675
Validation cohort	0.682 (0.746–0.811)	0.790 (0.687–0.893)	0.800	0.595	0.614	0.752 (0.677–0.826)	0.675	0.651	0.663
3D-PGRFM									
Primary cohort	0.753 (0.793–0.832)	0.781 (0.712–0.851)	0.763	0.688	0.696	0.819 (0.774–0.865)	0.799	0.717	0.761
Validation cohort	0.706 (0.767–0.827)	0.778 (0.672–0.884)	0.667	0.710	0.706	0.784 (0.713–0.855)	0.740	0.744	0.742
2D-PGRFM									
Primary cohort	0.714 (0.757–0.800)	0.778 (0.708–0.848)	0.842	0.656	0.678	0.773 (0.722–0.824)	0.753	0.691	0.724
Validation cohort	0.728 (0.784–0.840)	0.781 (0.681–0.880)	0.867	0.730	0.742	0.806 (0.739–0.872)	0.662	0.756	0.712
2D-PGRFM + Clinic									
Primary cohort	0.720 (0.763–0.805)	0.781 (0.712–0.850)	0.842	0.667	0.687	0.779 (0.729–0.829)	0.759	0.691	0.727
Validation cohort	0.731 (0.786–0.842)	0.783 (0.685–0.881)	0.867	0.730	0.742	0.809 (0.743–0.875)	0.649	0.744	0.699
ROSRM									
Primary cohort	0.737 (0.769–0.802)	0.786 (0.743–0.828)	0.855	0.645	0.715	0.807 (0.764–0.849)	0.806	0.704	0.772
Validation cohort	0.727 (0.768–0.810)	0.765 (0.706–0.824)	0.849	0.692	0.744	0.835 (0.783–0.886)	0.704	0.779	0.729

GS: Gleason score; AUC: area under the receiver operating characteristic (ROC) curve; SP: specificity; SE: sensitivity; CM: clinical model; PGRFM: prostate gland radiomic features based model; ROSRM: radiomic model constructed by features after the random oversampling.

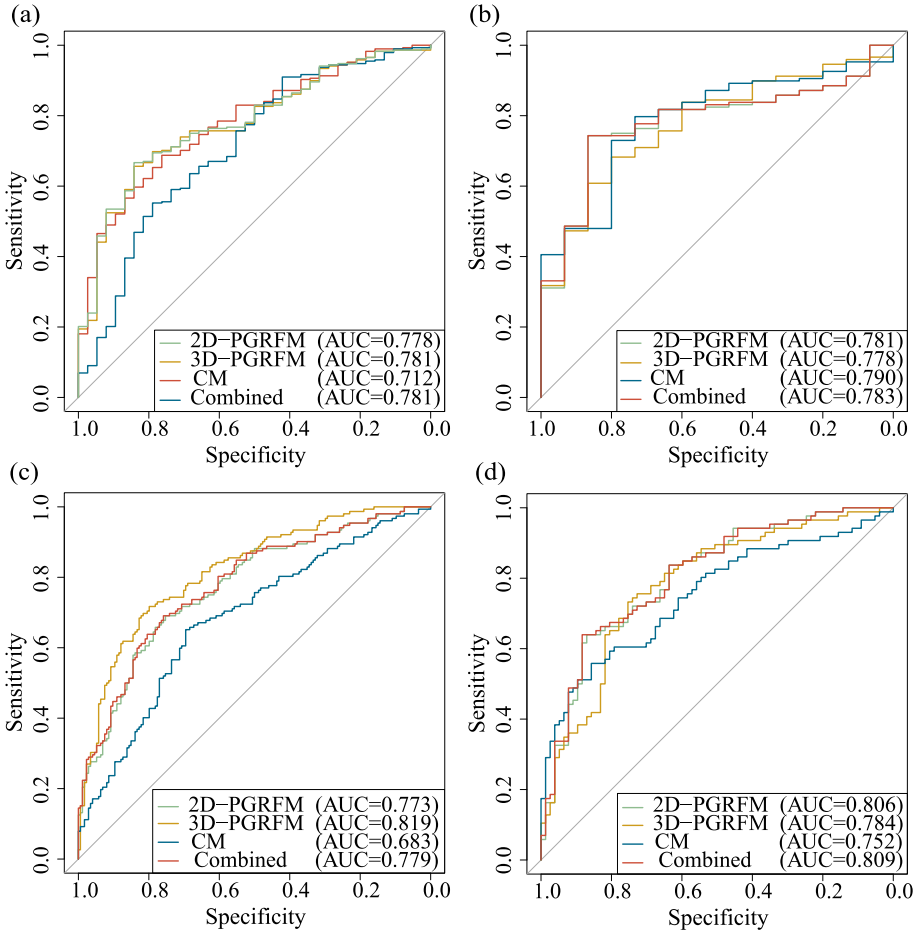

Fig. 3. ROC curves for predicting (a–b) GS < 7 vs. GS ⩾ 7 and (c–d) GS ⩽ 7 vs. GS > 7. (a) and (c) are in the primary cohort, and (b) and (d) are in the validation cohort. GS: Gleason score; CM: clinical model; PGRFM: radiomic GS prediction model; ROC: receiver operating characteristic.

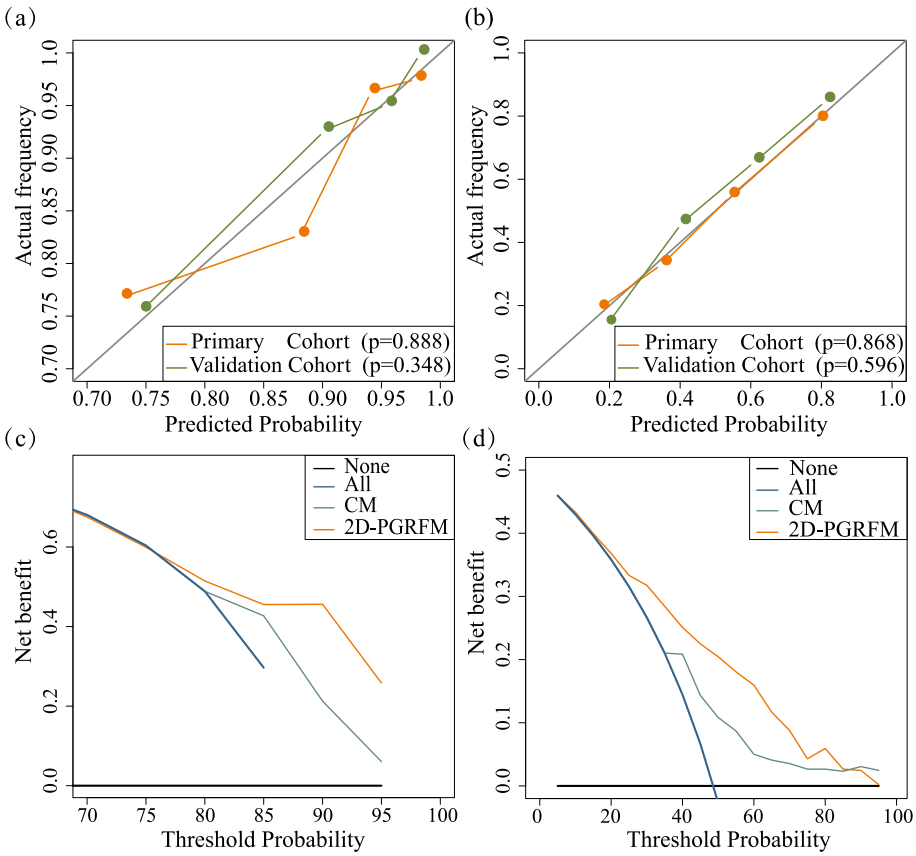

Fig. 4. Calibration curves and decision curves of PGRFM. (a–b) are calibration curves. (c–d) are decision curves. (a) and (c) are for GS < 7 vs. GS ⩾ 7, and (b) and (d) are for GS ⩽ 7 vs. GS > 7.

cohorts. In addition, whether for prediction of GS < 7 vs. GS ⩾ 7 or GS ⩽ 7 vs. GS > 7, the decision curves displayed relatively good performance for the PGRFM in clinical applications. Although the probability of achieving accurate prediction ranges from 0% to 100%, the decision curves of PGRFM showed a greater advantage than either the scheme in which all patients were assumed to achieve accurate prediction or the scheme in which no patients were.

In terms of a comprehensive evaluation, 2D-PGRFM got a relatively better performence than 3D-PGRFM, and we therefore carried out following the analyses with the 2D-PGRFM. The combined model constructed by the 2D radiomic key features and clinic information (2D-PGRFM + Clinic) got a little improved C-index than 2D-PGRFM in the validation cohort (2D-PGRFM + Clinic: 0.731 (0.786–0.842), 2D-PGRFM: 0.728 (0.784–0.840)), while the other results were similar to 2D-PGRFM.

3.4. Feature stability in random datasets

The detailed distributions of the key feature rankings for 2D-PGRFM are shown in Fig. 5. As shown in Fig. 5(a)–(c), we found that under 30 random datasets most of the key features of T2WI, ADC and DWI sequences had a raincloud plot distributed at a rank of less than 0.30, which indicated most of the key features can be selected ignoring the change of training data. This stability was more obvious when we comprehensively considered each sequence. As shown in Fig. 5(d), the key features of the T2WI, ADC, and DWI sequences had an average ranking of 6.065%, 7.485%, and 11.506% in the random datasets, respectively. That means the key features of all of the sequences of MRI ranked on average at the top 12% within other random primary cohorts, which demonstrated the satisfactory stability of the key radiomic features of the prostate gland in spite of the variation of the primary cohorts.

3.5. Tolerance for segmentation deviation

Four derived ROIs were generated by controlling the distance to the original manual segmentation: 1 mm intended ROI, 1 mm expanded ROI, 2 mm intended ROI, and 2 mm expanded ROI respectively. Accordingly, four derived radiomic models based on these four derived ROIs were constructed and validated. As shown in Table 3, the performance of the 4 derived radiomic models had a decline compared with 2D-PGRFM based on the original manual segmentation in the validation cohort, which was more obvious as the distance increased and more evident for the expanded ROI than the contracted ROI. However, it is worth noting that the performance decrease was not significant for the 1 mm intended ROI (GS < 7 vs. GS ⩾ 7: p = 0.732, GS ⩽ 7 vs. GS > 7: p = 0.157), 1 mm expanded ROI (GS < 7 vs. GS ⩾ 7: p = 0.441, GS ⩽ 7 vs. GS > 7: p = 0.407), and 2 mm intended ROI (GS < 7 vs. GS ⩾ 7: p = 0.854, GS ⩽ 7 vs. GS > 7: p = 0.146) except for the 2 mm expanded ROI in the prediction of GS ⩽ 7 vs. GS > 7 (p < 0.05).

3.6. Influence of unbalanced data distribution

After ROS, the cases number of GS < 7, GS = 7, and GS > 7 were oversampled from 38, 136, and 152 to 152, 152, and 152, respectively. For identifying GS < 7, except for a similar ACC (ROSRM: 0.744, 2D-PGRFM: 0.742), our 2D radiomic model based on ROS (ROSRM) had no improvement in other results in the validation cohort. However, for the prediction of GS > 7, ROSRM improved the performance of 2D-PGRFM in all of the AUC, SP, SE, and ACC in the validation cohort (ROSRM: AUC = 0.835; SP = 0.704, SE = 0.779, ACC = 0.729, 2D-PGRFM: AUC = 0.806; SP = 0.662, SE = 0.756, ACC = 0.712), as

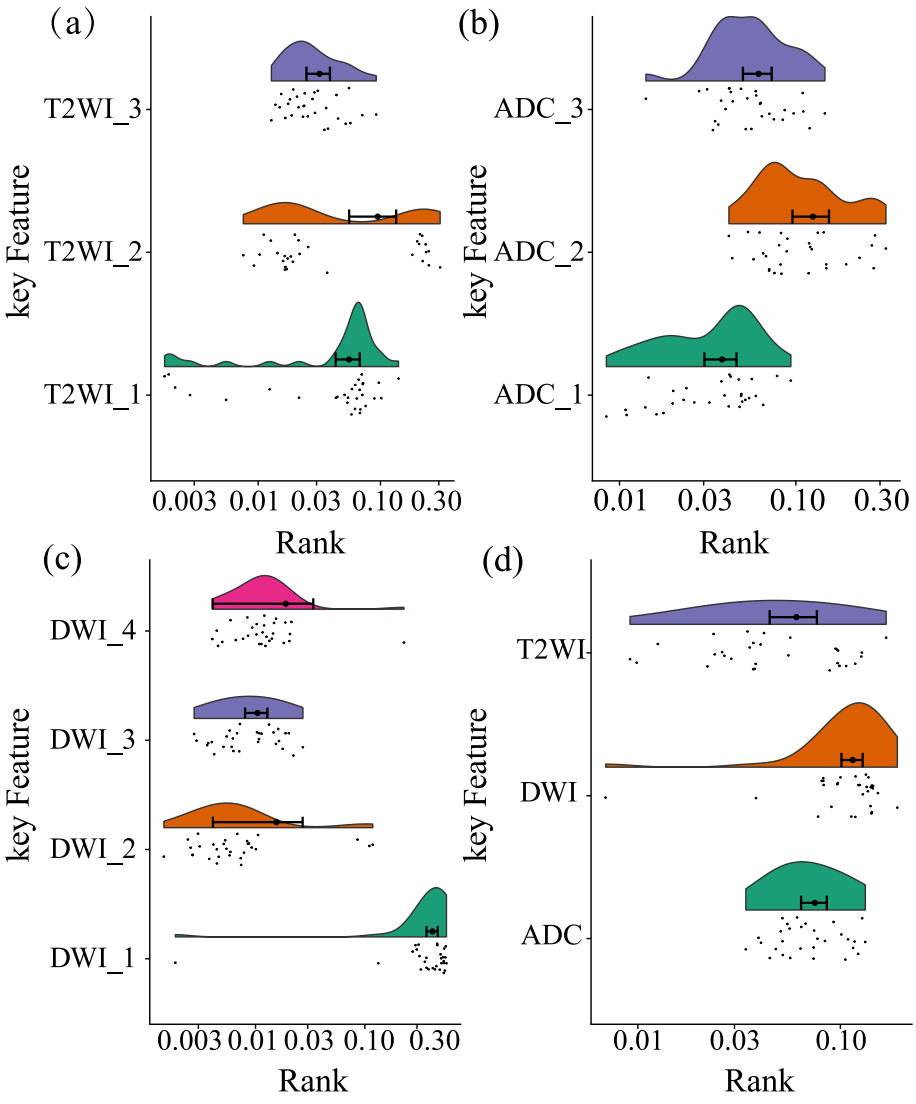

Fig. 5. The raincloud plot of the key radiomic features. The ranking of the key features of T2WI (a), ADC (b), and DWI (c), and the average ranking for three sequences (d). The vertical axis respectively represents the targeted key features of sequence T2WI, ADC, and DWI, while the horizontal axis shows the corresponding feature ranking.

shown in Table 2.

4. Discussion

PCa tends to be multi-lesioned and has a scattered distribution within the prostate gland. The decision-making about PCa depends mainly on the Gleason score from needle biopsy, which, however, has lesion location limitations and adverse effects. Compared to the local lesion, the prostate gland includes multi-lesion information and the peritumoral environment that has been reported to be helpful for tumor diagnosis or prognostic analysis [48–50]. Besides, VPG has been reported to be associated with PCa. Therefore, it is natural to attempt to explore further information on the prostate gland and discuss whether it is feasible to segment the whole prostate gland as the ROI for identifying GS.

Our study utilized a radiomics analysis on the prostate gland ROI and develop a non-invasive method for predicting GS < 7, GS $= 7$, and GS > 7 among PCa patients. The 3D and 2D-PGRFM models were constructed based on the 3D and 2D ROIs, respectively. Compared to clinical model, both of them obtained more satisfactory results with a C-index higher than 0.700 and an average AUC higher than 0.780, which indicated the values of the prostate gland radiomic features in predicting GS. Specifically, 2D-PGRFM acquired a higher C-index in predicting GS < 7 vs. GS $= 7$ vs. GS > 7 than 3D-PGRFM, and a possible reason for this may be due to the noise incurred by resampling the 3D ROIs thickness on the z-axis. This result was consistent with a former comparative study between the 2D and 3D CT radiomic features [36]. Within the 2D key features, textural features contributed the most, especially when computed by GLRLM. This prompted the heterogeneity of PCa, and the non-invasive diagnosis preference tended to the roughness of the texture.

Compared with identifying GS < 7, our 2D-PGRFM performed better at identifying GS > 7. The reason may be that there were more invaded regions in the prostate gland of patients with high-grade PCa. However, the specificity of identifying GS > 7 was decreased when 2D-PGRFM was combined with the clinical characteristics, probably because the clinical predictors were not significant for GS compared to the radiomic signatures.

We investigated the stability of the prostate gland as an ROI in predicting GS. On the one hand, we inspected the stability of the key radiomic features in different data cohorts. Our study showed that the key features involved in the 2D-PGRFM construction maintained a top importance ranking in the other 30 random primary cohorts. This indicated the key radiomic features of the prostate gland had good

Table 3
Model Performance on the derived ROIs.

Derived ROI	GS < 7 vs. GS = 7 vs. GS > 7	GS < 7 vs. GS ⩾ 7				GS ⩽ 7 vs. GS > 7			
	C-index	AUC	SP	SE	ACC	AUC	SP	SE	ACC
1 mm									
Indented ROI									
Primary cohort	0.744 (0.785–0.827)	0.794 (0.725–0.863)	0.868	0.677	0.699	0.803 (0.755–0.851)	0.937	0.533	0.749
Validation cohort	0.690 (0.751–0.812)	0.743 (0.634–0.851)	0.667	0.723	0.718	0.771 (0.698–0.843)	0.883	0.500	0.681
Expanded ROI									
Primary cohort	0.736 (0.776–0.816)	0.793 (0.732–0.854)	0.868	0.663	0.687	0.794 (0.745–0.842)	0.695	0.790	0.739
Validation cohort	0.711 (0.767–0.823)	0.720 (0.625–0.815)	0.733	0.676	0.681	0.799 (0.732–0.866)	0.636	0.767	0.706
2 mm									
Indented ROI									
Primary cohort	0.746 (0.787–0.828)	0.767 (0.691–0.843)	0.737	0.729	0.730	0.815 (0.769–0.860)	0.839	0.665	0.758
Validation cohort	0.690 (0.751–0.812)	0.753 (0.649–0.857)	0.667	0.737	0.730	0.769 (0.697–0.841)	0.727	0.674	0.699
Expanded ROI									
Primary cohort	0.725 (0.766–0.807)	0.766 (0.699–0.833)	0.921	0.545	0.589	0.791 (0.769–0.860)	0.707	0.763	0.733
Validation cohort	0.646 (0.708–0.769)	0.705 (0.604–0.805)	0.800	0.547	0.571	0.727* (0.650–0.804)	0.584	0.640	0.614

ROI: region of interest; GS: Gleason score; CI: confidence interval; AUC: area under the receiver operating characteristic (ROC) curve; SP: specificity; SE: sensitivity; *: There was a significant difference between the target and 2D-PGRFM.

stability being free from the influence of the random data cohort when predicting GS. On the other hand, we introduced a distance matrix based algorithm that could generate derived ROIs with different sizes to simulate the manual segmentation error. We found no significant performance degradation as the distance to the prostate gland boundary increased by 1 mm. However, the results showed that adding more external surroundings to the prostate gland, such as the fat or muscle, would induce apparent noise to the radiomic model. This indicated that the prostate gland segmentation can tolerate an error of up to 2 mm while maintaining the stable performance. In addition, our radiomic features were less sensitive to the null value, outliers, and MRI resolution, due to the data pre-process and interpolation.

Certain studies investigated the radiomic method to identify GS. Fehr et al. [17] presented multiple machine learning methods for classifying GS = 6 vs. GS ⩾ 7, focusing on textural features and two-subgroup classification. Wibmer et al. [51] investigated the roles of Haralick texture from T2W MRI in differentiating GS = 6 vs. GS = 7, GS = 6 vs. GS > 7, and GS ⩽ 7 vs. GS > 7. Chaddad et al. [52] successfully predict GS ⩽ 7, GS = 7, and GS ⩾ 7 using mpMRI texture features, respectively. However, they did not focus on simultaneously predicting these groups. Moreover, previous studies mainly analyzed the information of tumor regions, involving sophisticated tumor annotations, and paid little attention to VPG. Tanaka et al. [29] found VPG showed a significantly more vital predictive ability than serum PSA for identifying PCa. Consistently, Peng et al. [28] claimed that the VPG had similar predictive performance compared to central-gland volumes from T2WI for the diagnosis of PCa. In our study, although the VPG exhibited a strong relation to the GS in the primary cohort, it was found to have a weak correlation to the GS in the validation cohort. This may be a result of a certain difference between the radiomic volume and the actual volume and needs to be examined further in a future study involving the accurately measured volume. But, our study demonstrated the potential of prostate gland radiomic features in achieving three-way identification of GS < 7 vs. GS = 7 vs. GS > 7. This is valuable for recommended treatment strategies in clinics. Patients with GS < 7 should be monitored closely, while those with GS > 7 should undergo more aggressive symptomatic treatment. Given that GS = 7 represents a relatively specific group, further examination is recommended in this population.

The contributing results of prostate-gland-ROI may mainly be due to comprehensive information, including the scattered lesion itself and its surroundings, reflecting tissue microcirculation and reducing the missed diagnosis by capturing occult lesions that are often omitted by naked eyes. Also, since the features we used were extracted from the whole prostate gland rather than the lesions, we need to worry little about the artificial noise to the tumor volume. Meanwhile, the prostate gland segmentation was less influenced by the mismatching led by image deformation between T2WI and DWI sequences. In addition, compared to the refined and complex tumour lesion segmentation, simplified and rougher gland segmentation is helpful to achieve auto segmentation in the future, which is valuable for the realization and application of the fully automated radiomic analysis. Therefore, PGRFM development based on prostate gland radiomic features may aid in clinical decision-making and promote the automated non-invasive prediction of GS.

Considering the imbalance of the class occupation could have an impact on the performance of the model, we applied ROS and found that the performance of 2D-PGRFM was improved, especially the specificity in predicting GS > 7, which indicated a space for performance improvement of our 2D-PGRFM in a larger dataset.

The present study possesses some limitations. Firstly, the radiomic model constructed based on lesion segmentation was not involved and needs to be further compared with our model, which will be investigated in future studies. Second, it would be more rigorous to involve multiple raters to guarantee the accurate segmentation. Besides, as this was a single-center study, further studies involving multiple centers are required to determine the generalizability of our model. Lastly, only one b-value of 800 s/mm^2 was utilized for DWI in our study, and more b-value series are supposed to be discussed in further studies.

5. Conclusion

This study explored the predictive potential of the prostate gland radiomic features in identifying GS, and evaluated the stability of the prostate gland as an ROI from the aspects of key feature stability and segmentation boundary error tolerance, respectively. Firstly, the satisfactory results of radiomic models, especially the 2D radiomic model, indicated the potential of the prostate gland radiomic features in

predicting GS. Secondly, our study proved that the prostate gland exhibited good stability as an ROI in predicting GS. For one thing, our study indicated the key radiomic features of the prostate gland had good stability being free from the influence of the random data cohort when predicting GS. For another, according to our distance matrix based algorithm, the result indicated that the prostate gland segmentation could tolerate an error of up to 2 mm maintaining the stable performance. In summary, the results of our present study demonstrated that a prostate-gland-radiomics-based model incorporating high-dimensional MRI features has the potential to be used as an assistance for the non-invasive, accurate prediction of preoperative GS in PCa patients.

Declaration of competing interest

There is no conflict of interest to declare.

Acknowledgments

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB 38040200), National Natural Science Foundation of China (82022036, 91959130, 81971776, 81771924, 62027901, 81930053), National Key R&D Program of China (2017YFA 0205200), the Beijing Natural Science Foundation (L182061, Z20J00105), Chinese Academy of Sciences under Grant No. GJJSTD2017 0004 and QYZDJ-SSW-JSC005, the Project of High-Level Talents Team Introduction in Zhuhai City (Zhuhai HLHPTP201703), and the Youth Innovation Promotion Association CAS (2017175). The authors would like to acknowledge the instrumental and technical support of Multi-modal biomedical imaging experimental platform, Institute of Automation, Chinese Academy of Sciences.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2022.105318.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, CA A Cancer J. Clin. 69 (1) (2019) 7–34, https://doi.org/10.3322/caac.21551, 2019.
[2] S. Negoita, E.J. Feuer, A. Mariotto, K.A. Cronin, V.I. Petkov, S.K. Hussey, V. Benard, S.J. Henley, R.N. Anderson, S. Fedewa, et al., Annual report to the nation on the status of cancer, part ii: recent changes in prostate cancer trends and disease characteristics, Cancer 124 (13) (2018) 2801–2814, https://doi.org/10.1002/cncr.31549.
[3] V.A. Moyer, Screening for prostate cancer: us preventive services task force recommendation statement, Ann. Intern. Med. 157 (2) (2012) 120–134, https://doi.org/10.7326/0003-4819-157-2-201207170-00459.
[4] S.A. Fedewa, E.M. Ward, O. Brawley, A. Jemal, Recent patterns of prostate-specific antigen testing for prostate cancer screening in the United States, JAMA Intern. Med. 177 (7) (2017) 1040–1042, https://doi.org/10.1001/jamainternmed.2017.0340.
[5] P.H. Carroll, J.L. Mohler, Nccn guidelines updates: prostate cancer and prostate cancer early detection, J. Natl. Compr. Cancer Netw. 16 (5S) (2018) 620–623, https://doi.org/10.6004/jnccn.2018.0036.
[6] J.I. Epstein, L. Egevad, M.B. Amin, B. Delahunt, J.R. Srigley, P.A. Humphrey, The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma, Am. J. Surg. Pathol. 40 (2) (2016) 244–252, https://doi.org/10.1097/PAS.0000000000000530.
[7] O. Hassan, M. Han, A. Zhou, A. Paulk, Y. Sun, A. Al-Harbi, A. Alrajjal, F.B. Dos Santos, J.I. Epstein, Incidence of extraprostatic extension at radical prostatectomy with pure gleason score 3+ 3= 6 (grade group 1) cancer: implications for whether gleason score 6 prostate cancer should be renamed" not cancer" and for selection criteria for active surveillance, J. Urol. 199 (6) (2018) 1482–1487, https://doi.org/10.1016/j.juro.2017.11.067.
[8] M.G. Sanda, J.A. Cadeddu, E. Kirkby, R.C. Chen, T. Crispino, J. Fontanarosa, S. J. Freedland, K. Greene, L.H. Klotz, D.V. Makarov, et al., Clinically localized prostate cancer: aua/astro/suo guideline. part i: risk stratification, shared decision making, and care options, J. Urol. 199 (3) (2018) 683–690, https://doi.org/10.1016/j.juro.2017.11.095.
[9] H. Miyake, I. Sakai, T.-a. Inoue, I. Hara, M. Fujisawa, The limited significance of a longer duration of neoadjuvant hormonal therapy prior to radical prostatectomy for high-risk prostate cancer in Japanese men, Urol. Int. 77 (2) (2006) 122–126, https://doi.org/10.1159/000093904.
[10] J. Xiang, H. Yan, J. Li, X. Wang, H. Chen, X. Zheng, Transperineal versus transrectal prostate biopsy in the diagnosis of prostate cancer: a systematic review and meta-analysis, World J. Surg. Oncol. 17 (1) (2019) 1–11, https://doi.org/10.1186/s12957-019-1573-0.
[11] M. Seles, T. Gutschi, K. Mayrhofer, K. Fischereder, G. Ehrlich, G. Gallé, S. Gutschi, O. Pachernegg, K. Pummer, H. Augustin, Sampling of the anterior apical region results in increased cancer detection and upgrading in transrectal repeat saturation biopsy of the prostate, BJU Int. 117 (4) (2016) 592–597, https://doi.org/10.1111/bju.13108.
[12] M.A. Liss, B. Ehdaie, S. Loeb, M.V. Meng, J.D. Raman, V. Spears, S.P. Stroup, An update of the american urological association white paper on the prevention and treatment of the more common complications related to prostate biopsy, J. Urol. 198 (2) (2017) 329–334, https://doi.org/10.1016/j.juro.2017.01.103.
[13] A.R. Meyer, M. Mamawala, J.S. Winoker, P. Landis, J.I. Epstein, K.J. Macura, M. E. Allaf, A.W. Partin, C.P. Pavlovich, M.A. Gorin, Transperineal prostate biopsy improves the detection of clinically significant prostate cancer among men on active surveillance, J. Urol. 205 (4) (2021) 1069–1074, https://doi.org/10.1097/JU.0000000000001523.
[14] B.T. Ristau, M. Allaway, D. Cendo, J. Hart, J. Riley, V. Parousis, P.C. Albertsen, Free-hand transperineal prostate biopsy provides acceptable cancer detection and minimizes risk of infection: evolving experience with a 10-sector template, in: Urologic Oncology: Seminars and Original Investigations, vol. 36, Elsevier, 2018, pp. 528–e15, https://doi.org/10.1016/j.urolonc.2018.09.013.
[15] D.T. Chang, B. Challacombe, N. Lawrentschuk, Transperineal biopsy of the prostate—is this the future? Nat. Rev. Urol. 10 (12) (2013) 690–702, https://doi.org/10.1038/nrurol.2013.195.
[16] L. Dickinson, H.U. Ahmed, C. Allen, J.O. Barentsz, B. Carey, J.J. Futterer, S. W. Heijmink, P.J. Hoskin, A. Kirkham, A.R. Padhani, et al., Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a european consensus meeting, Eur. Urol. 59 (4) (2011) 477–494, https://doi.org/10.1016/j.eururo.2010.12.009.
[17] D. Fehr, H. Veeraraghavan, A. Wibmer, T. Gondo, K. Matsumoto, H.A. Vargas, E. Sala, H. Hricak, J.O. Deasy, Automatic classification of prostate cancer gleason scores from multiparametric magnetic resonance images, Proc. Natl. Acad. Sci. Unit. States Am. 112 (46) (2015) E6265–E6273, https://doi.org/10.1073/pnas.1505935112.
[18] G. Nketiah, M. Elschot, E. Kim, J.R. Teruel, T.W. Scheenen, T.F. Bathen, K. M. Selnæs, T2-weighted mri-derived textural features reflect prostate cancer aggressiveness: preliminary results, Eur. Radiol. 27 (7) (2017) 3050–3059, https://doi.org/10.1007/s00330-016-4663-1.
[19] J.C. Weinreb, J.O. Barentsz, P.L. Choyke, F. Cornud, M.A. Haider, K.J. Macura, D. Margolis, M.D. Schnall, F. Shtern, C.M. Tempany, et al., Pi-rads prostate imaging-reporting and data system: 2015, version 2, Eur. Urol. 69 (1) (2016) 16–40, https://doi.org/10.1016/j.eururo.2015.08.052.
[20] Y. Ueno, T. Tamada, V. Bist, C. Reinhold, H. Miyake, U. Tanaka, K. Kitajima, K. Sugimura, S. Takahashi, Multiparametric magnetic resonance imaging: current role in prostate cancer management, Int. J. Urol. 23 (7) (2016) 550–557, https://doi.org/10.1111/iju.13119.
[21] B. Turkbey, A.M. Brown, S. Sankineni, B.J. Wood, P.A. Pinto, P.L. Choyke, Multiparametric prostate magnetic resonance imaging in the evaluation of prostate cancer, CA A Cancer J. Clin. 66 (4) (2016) 326–336, https://doi.org/10.3322/caac.21333.
[22] N. Mottet, J. Bellmunt, M. Bolla, E. Briers, M.G. Cumberbatch, M. De Santis, N. Fossati, T. Gross, A.M. Henry, S. Joniau, et al., Eau-estro-siog guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent, Eur. Urol. 71 (4) (2017) 618–629, https://doi.org/10.1016/j.eururo.2016.08.003.
[23] T.B. Lam, S. MacLennan, P.-P.M. Willemse, M.D. Mason, K. Plass, R. Shepherd, R. Baanders, C.H. Bangma, A. Bjartell, A. Bossi, et al., Eau-eanm-estro-esur-siog prostate cancer guideline panel consensus statements for deferred treatment with curative intent for localised prostate cancer from an international collaborative study (detective study), Eur. Urol. 76 (6) (2019) 790–813, https://doi.org/10.1016/j.eururo.2019.09.020.
[24] C.K. Kuhl, R. Bruhn, N. Krämer, S. Nebelung, A. Heidenreich, S. Schrading, Abbreviated biparametric prostate mr imaging in men with elevated prostate-specific antigen, Radiology 285 (2) (2017) 493–505, https://doi.org/10.1148/radiol.2017170129.
[25] E. Di Campli, A.D. Pizzi, B. Seccia, R. Cianci, M. d'Annibale, A. Colasante, S. Cinalli, P. Castellan, R. Navarra, R. Iantorno, et al., Diagnostic accuracy of biparametric vs multiparametric mri in clinically significant prostate cancer: comparison between readers with different experience, Eur. J. Radiol. 101 (2018) 17–23, https://doi.org/10.1016/j.ejrad.2018.01.028.
[26] H.S. Sidhu, S. Benigno, B. Ganeshan, N. Dikaios, E.W. Johnston, C. Allen, A. Kirkham, A.M. Groves, H.U. Ahmed, M. Emberton, et al., Textural analysis of multiparametric mri detects transition zone prostate cancer, Eur. Radiol. 27 (6) (2017) 2348–2358, https://doi.org/10.1007/s00330-016-4579-9.
[27] L. Gong, M. Xu, M. Fang, J. Zou, S. Yang, X. Yu, D. Xu, L. Zhou, H. Li, B. He, et al., Noninvasive prediction of high-grade prostate cancer via biparametric mri radiomics, J. Magn. Reson. Imag. 52 (4) (2020) 1102–1109, https://doi.org/10.1002/jmri.27132.
[28] Y. Peng, D. Shen, S. Liao, B. Turkbey, S. Rais-Bahrami, B. Wood, I. Karademir, T. Antic, A. Yousef, Y. Jiang, et al., Mri-based prostate volume-adjusted prostate-specific antigen in the diagnosis of prostate cancer, J. Magn. Reson. Imag. 42 (6) (2015) 1733–1739, https://doi.org/10.1002/jmri.24944.
[29] I. Karademir, D. Shen, Y. Peng, S. Liao, Y. Jiang, A. Yousuf, G. Karczmar, S. Sammet, S. Wang, M. Medved, et al., Prostate volumes derived from mri and

volume-adjusted serum prostate-specific antigen: correlation with gleason score of prostate cancer, AJR Am. J. Roentgenol. 201 (5) (2013) 1041–1048, https://doi.org/10.2214/AJR.13.10591.
[30] S.P. Krafft, A. Rao, F. Stingo, T.M. Briere, L.E. Court, Z. Liao, M.K. Martel, The utility of quantitative ct radiomics features for improved prediction of radiation pneumonitis, Med. Phys. 45 (11) (2018) 5317–5324, https://doi.org/10.1002/mp.13150.
[31] F. Homayounieh, S. Ebrahimian, R. Babaei, H.K. Mobin, E. Zhang, B.C. Bizzo, I. Mohseni, S.R. Digumarthy, M.K. Kalra, Ct radiomics, radiologists, and clinical information in predicting outcome of patients with covid-19 pneumonia, Radiology: Cardiothoracic Imaging 2 (4) (2020), e200322, https://doi.org/10.1148/ryct.2020200322.
[32] B. Turkbey, A.B. Rosenkrantz, M.A. Haider, A.R. Padhani, G. Villeirs, K.J. Macura, C.M. Tempany, P.L. Choyke, F. Cornud, D.J. Margolis, et al., Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2, Eur. Urol. 76 (3) (2019) 340–351, https://doi.org/10.1016/j.eururo.2019.02.033.
[33] R. Etzioni, A. Tsodikov, A. Mariotto, A. Szabo, S. Falcon, J. Wegelin, K. Karnofski, R. Gulati, D.F. Penson, E. Feuer, et al., Quantifying the role of psa screening in the us prostate cancer mortality decline, Cancer Causes Control 19 (2) (2008) 175–181, https://doi.org/10.1007/s10552-007-9083-8.
[34] C.L. Vendrami, R.J. McCarthy, A. Chatterjee, D. Casalino, E.M. Schaeffer, W. J. Catalona, F.H. Miller, The utility of prostate specific antigen density, prostate health index, and prostate health index density in predicting positive prostate biopsy outcome is dependent on the prostate biopsy methods, Urology 129 (2019) 153–159, https://doi.org/10.1016/j.urology.2019.03.018.
[35] S.M. Bruinsma, C.H. Bangma, P.R. Carroll, M.S. Leapman, A. Rannikko, N. Petrides, M. Weerakoon, L.P. Bokhorst, M.J. Roobol, B. Ehdaie, et al., Active surveillance for prostate cancer: a narrative review of clinical guidelines, Nat. Rev. Urol. 13 (3) (2016) 151–167, https://doi.org/10.1038/nrurol.2015.313.
[36] L. Meng, D. Dong, X. Chen, M. Fang, R. Wang, J. Li, Z. Liu, J. Tian, 2d and 3d ct radiomic features performance comparison in characterization of gastric cancer: a multi-center study, IEEE.J. Biomed. Health Inf. 25 (3) (2020) 755–763, https://doi.org/10.1109/JBHI.2020.3002805.
[37] J.J. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H.J. Aerts, Computational radiomics system to decode the radiographic phenotype, Cancer Res. 77 (21) (2017) e104–e107, https://doi.org/10.1158/0008-5472.CAN-17-0339.
[38] D. Bonekamp, S. Kohl, M. Wiesenfarth, P. Schelb, J.P. Radtke, M. Götz, P. Kickingereder, K. Yaqubi, B. Hitthaler, N. Gählert, et al., Radiomic machine learning for characterization of prostate lesions with mri: comparison to adc values, Radiology 289 (1) (2018) 128–137, https://doi.org/10.1148/radiol.2018173064.
[39] A. Algohary, S. Viswanath, R. Shiradkar, S. Ghose, S. Pahwa, D. Moses, I. Jambor, R. Shnier, M. Böhm, A.-M. Haynes, et al., Radiomic features on mri enable risk categorization of prostate cancer patients on active surveillance: preliminary findings, J. Magn. Reson. Imag. 48 (3) (2018) 818–828, https://doi.org/10.1002/jmri.25983.
[40] S.E. Viswanath, P.V. Chirra, M.C. Yim, N.M. Rofsky, A.S. Purysko, M.A. Rosen, B. N. Bloch, A. Madabhushi, Comparing radiomic classifiers and classifier ensembles for detection of peripheral zone prostate tumors on t2-weighted mri: a multi-site study, BMC Med. Imag. 19 (1) (2019) 1–12, https://doi.org/10.1186/s12880-019-0308-6.
[41] X. Min, M. Li, D. Dong, Z. Feng, P. Zhang, Z. Ke, H. You, F. Han, H. Ma, J. Tian, et al., Multi-parametric mri-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: cross-validation of a machine learning method, Eur. J. Radiol. 115 (2019) 16–21, https://doi.org/10.1016/j.ejrad.2019.03.010.
[42] E. Theodorsson-Norheim, Kruskal-wallis test: basic computer program to perform nonparametric one-way analysis of variance and multiple comparisons on ranks of several independent samples, Comput. Methods Progr. Biomed. 23 (1) (1986) 57–62, https://doi.org/10.1016/0169-2607(86)90081-7.
[43] H. Peng, F. Long, C. Ding, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, IEEE Trans. Pattern Anal. Mach. Intell. 27 (8) (2005) 1226–1238, https://doi.org/10.1109/TPAMI.2005.159.
[44] K.Z. Mao, Orthogonal forward selection and backward elimination algorithms for feature subset selection, IEEE. Trans. Syst. Man Cybern. B (Cybern.) 34 (1) (2004) 629–634, https://doi.org/10.1109/TSMCB.2002.804363.
[45] M.B. Shirzad, M.R. Keyvanpour, A feature selection method based on minimum redundancy maximum relevance for learning to rank, in: 2015 AI & Robotics (IRANOPEN), IEEE, 2015, pp. 1–5, https://doi.org/10.1109/RIOS.2015.7270735.
[46] F.E. Harrell Jr., Regression Modeling Strategies: with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Springer, 2015.
[47] M. Allen, D. Poggiali, K. Whitaker, T. R. Marshall, J. van Langen, R. A. Kievit, Raincloud plots: a multi-platform tool for robust data visualization, Wellcome Open Resdoi:https://doi.org/10.12688/wellcomeopenres.15191.1.
[48] N. Beig, M. Khorrami, M. Alilou, P. Prasanna, N. Braman, M. Orooji, S. Rakshit, K. Bera, P. Rajiah, J. Ginsberg, et al., Perinodular and intranodular radiomic features on lung ct images distinguish adenocarcinomas from granulomas, Radiology 290 (3) (2019) 783–792, https://doi.org/10.1148/radiol.2018180910.
[49] N. Braman, P. Prasanna, J. Whitney, S. Singh, N. Beig, M. Etesami, D.D. Bates, K. Gallagher, B.N. Bloch, M. Vulchi, et al., Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for her2 (erbb2)-positive breast cancer, JAMA Netw. Open 2 (4) (2019), https://doi.org/10.1001/jamanetworkopen.2019.2561 e192561-e192561.
[50] X. Wu, D. Dong, L. Zhang, M. Fang, Y. Zhu, B. He, Z. Ye, M. Zhang, S. Zhang, J. Tian, Exploring the predictive value of additional peritumoral regions based on deep learning and radiomics: a multicenter study, Med. Phys. 48 (5) (2021) 2374–2385, https://doi.org/10.1002/mp.14767.
[51] A. Wibmer, H. Hricak, T. Gondo, K. Matsumoto, H. Veeraraghavan, D. Fehr, J. Zheng, D. Goldman, C. Moskowitz, S.W. Fine, et al., Haralick texture analysis of prostate mri: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different gleason scores, Eur. Radiol. 25 (10) (2015) 2840–2850, https://doi.org/10.1007/s00330-015-3701-8.
[52] A. Chaddad, M.J. Kucharczyk, T. Niazi, Multimodal radiomic features for the predicting gleason score of prostate cancer, Cancers 10 (8) (2018) 249, https://doi.org/10.3390/cancers10080249.