Original Article

Diagnostic efficacy and interobserver agreement among readers with variable experience of the Prostate Imaging for Recurrence Reporting system with whole-mount histology after androgen deprivation therapy as a reference

Zhangzhe Chen[1,2#], Bingni Zhou[1,3#], Wei Liu[1,3], Hualei Gan[3,4], Ruchuan Chen[1,3], Lirui Yang[3,4], Liangping Zhou[1,3], Xiaohang Liu[1,3]

[1]Department of Radiology, Fudan University Shanghai Cancer Center, Shanghai, China; [2]Department of Radiology, Shanghai Geriatric Medical Center, Shanghai, China; [3]Department of Oncology, Shanghai Medical College of Fudan University, Shanghai, China; [4]Department of Pathology, Fudan University Shanghai Cancer Center, Shanghai, China

Contributions: (I) Conception and design: X Liu, L Zhou; (II) Administrative support: X Liu, L Zhou; (III) Provision of study materials or patients: Z Chen, H Gan; (IV) Collection and assembly of data: Z Chen, B Zhou, W Liu, R Chen, L Yang; (V) Data analysis and interpretation: Z Chen, B Zhou, W Liu, R Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

Correspondence to: Xiaohang Liu, MD; Liangping Zhou, MD. Department of Oncology, Shanghai Medical College of Fudan University, Shanghai, China; Department of Radiology, Fudan University Shanghai Cancer Center, 270 Dong-An Road, Shanghai 200032, China. Email: liuxiang_1940@163.com; zhoulp_2022@163.com.

Background: The Prostate Imaging for Recurrence Reporting (PI-RR) system was recently proposed to assess the local recurrence of prostate cancer (PCa), but its exact performance for the prostate after radiotherapy or radical prostatectomy is difficult to determine. We aimed to evaluate the diagnostic performance and interreader agreement of this system using whole-mount histology of the prostate after androgen deprivation therapy (ADT) as the standard of reference.

Methods: In total, 119 patients with PCa post-ADT underwent multiparametric magnetic resonance imaging (mp-MRI) before prostatectomy. Three radiologists analyzed the MRI images independently, scoring imaging findings according to PI-RR. Spearman correlation was performed to assess the relationship between the percentage of sectors with residual cancer and PI-RR score. The diagnostic performance for detection of residual cancer was assessed on a per-sector basis. The chi-squared test was used to compare the cancer detection rate (CDR) among readers. Overall and pairwise interreader agreement in assigning PI-RR categories and residual cancer sectors with a score ≥3 or ≥4 were evaluated with the Cohen kappa coefficient.

Results: Histology revealed 209 sectors with residual cancer. The percentage of pathologically positive sectors increased with the increase in PI-RR score for all readers. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at a cutoff of score 3 ranged from 74.2% to 83.7%, 86.4% to 92.7%, 51.3% to 64.3%, and 95.4% to 96.9%, respectively, and at a cutoff of score 4, they ranged from 47.4% to 56.5%, 97.9% to 98.6%, 82.5% to 85.3%, and 91.6% to 92.9%, respectively. There was no significant difference among the CDR of readers. In PI-RR categories and detection of residual cancer sectors, overall interreader agreement was moderate for all readers, but agreement was higher between the more experienced readers (moderate to substantial) than between the more and less experienced readers (fair to moderate).

Conclusions: MRI scoring with the PI-RR assessment provided accurate evaluation of PCa after ADT, but readers' experience influenced interreader agreement and cancer diagnosis.

Keywords: Androgen deprivation therapy (ADT); diffusion-weighted imaging (DWI); magnetic resonance imaging (MRI); prostate cancer (PCa)

Submitted Nov 19, 2023. Accepted for publication Mar 06, 2024. Published online Mar 28, 2024.
doi: 10.21037/qims-23-1643
View this article at: https://dx.doi.org/10.21037/qims-23-1643

Introduction

The Prostate Imaging Reporting and Data System (PI-RADS) has been developed for treatment-naive patients, but its use in the posttreatment setting, especially after radiotherapy, has not been established. There is thus a need for well-designed and controlled means to evaluating patients with prostate cancer (PCa) using a standardized prostate magnetic resonance imaging (MRI) readout method (1).

The Prostate Imaging for Recurrence Reporting (PI-RR) was designed to promote standardization and decrease variation in the acquisition, interpretation, and reporting of MRI for assessing the local relapse of PCa, as well as to better guide evaluation following therapy (2). A simplified and standardized terminology for the content of the reports, in which five assessment categories are used to summarize the suspicion of local relapse, has been developed. A recent study proved that PI-RR could provide reproducible, structured, and accurate assessment for local recurrence in patients with PCa after definitive therapy. However, in this study, PI-RR assessment categories were assessed on a per-patient basis, and all readers were experienced radiologists (3). The actual recurrence rates in individual PI-RR categories based on lesion or sector, which could facilitate a more approximate assessment following biopsy or therapy, are currently unknown, and the effect of readers' experience on the efficacy PI-RR is also unclear.

Multiparametric magnetic resonance imaging (mp-MRI) for detection of relapse after radical radiotherapy (RT) has been examined in previous studies, but these studies were mostly based on the follow-up of patients undergoing needle biopsy, which is likely to miss some lesions and influence the assessment (4-11). A larger sample consisting of patients who have undergone whole-mount pathology, correlation with other imaging modalities, and clinical validation are therefore needed. However, the number of patients who have undergone radical prostatectomy (RP) after RT is rare, with the only two studies based on RP both using very small samples (12,13).

Androgen deprivation therapy (ADT), a vital form of neoadjuvant therapy before radiotherapy and surgery, is a key therapeutic approach for locally advanced and metastatic PCa (14). Prostate pathology after ADT shares a great deal in common with that after RT, with both including, for example, squamous metaplasia, marked atrophy, or a decrease in the number and reduction in the size of normal glandular acini (15,16), and both share similar MRI features (17-20). Moreover, many patients undergo RP after ADT, so a larger number of whole-mount pathology samples could be available. Therefore, an analogy can be made between patients post-ADT and those post-RT for the validation of the PI-RR system.

The purpose of this study was thus to primarily evaluate the interreader agreement and diagnostic ability of PI-RR using post-ADT prostate MRI images based on whole-mount pathology and to provide a reference for the further use of PI-RR in patients treated with RT. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-23-1643/rc).

Methods

Study population

This retrospective, single-center study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the Ethics Committee of Fudan University Shanghai Cancer Center. Informed consent was obtained from all individual participants.

From March 2017 to June 2022, 147 patients who underwent RP after neoadjuvant ADT and presurgical MRI examinations were retrieved from our picture archiving and communication system according to the following criteria

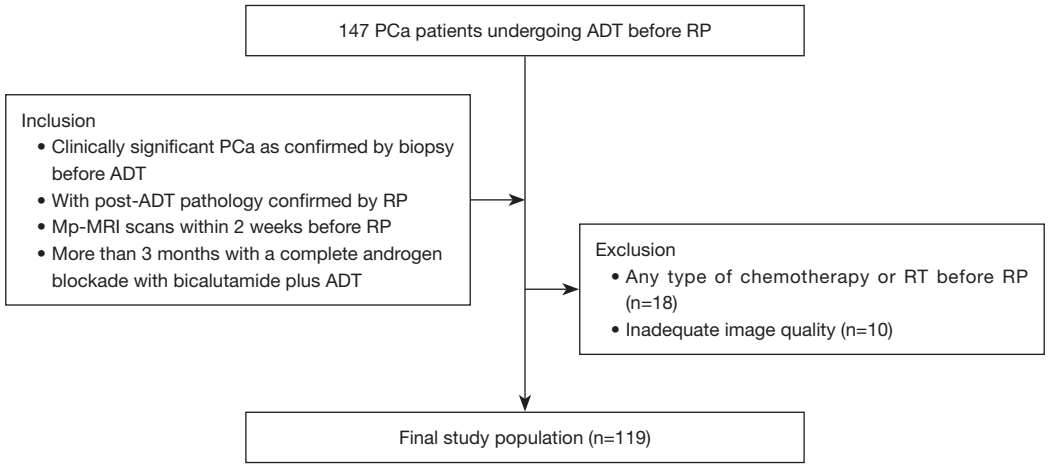

Figure 1 Flowchart of participant inclusion. PCa, prostate cancer; ADT, androgen deprivation therapy; RP, radical prostatectomy; RT, radical radiotherapy; mp-MRI, multiparametric magnetic resonance imaging.

(*Figure 1*).

(I) Patients had clinically significant PCa (Gleason score ≥7, greatest percentage of cancer >50% and more than two positive cores) confirmed by biopsy before ADT, with post-ADT pathology confirmed by RP. (II) Patients underwent post-ADT mp-MRI scans [T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE) imaging] in Fudan University Shanghai Cancer Center within 2 weeks before surgery, with the quality of images being sufficiently high for diagnosis and analysis. (III) Patients were treated for more than 3 months with a complete androgen blockade with bicalutamide plus ADT with goserelin, leuprolide, or abiraterone and did not receive other therapy.

Eighteen patients were excluded for any type of chemotherapy or RT before RP and 10 for poor-quality MRI images. Age, prostate specific antigen (PSA), duration of ADT, Gleason score and M stage were collected from each included patient.

MRI

Preoperative MRI was performed on two 3.0-T MRI systems (MAGNETOM Skyra 3.0 T, Siemens Healthineers, Erlangen, Germany). The entire prostate gland and seminal vesicles were imaged on coronal, sagittal, and axial slices using T2WI, DWI, and DCE. T2WI was completed with a fast-recovery fast-spin-echo (FR-FSE) sequence [repetition time (TR)/echo time (TE), 7,120 ms/89 ms; number of excitations, 2; slice thickness, 3 mm; spacing, 1 mm; matrix 324×320]. T1WI was completed with a fast spoiled gradient-echo (FSPGR) sequence (TR/TE, 231 ms/2.5 ms; slice thickness, 5.5 mm; spacing, 1 mm; matrix 204×320). DWI was completed with a readout-segmented echo-planar imaging (RS-EPI)-DWI sequence (TR/TE, 4,670 ms/63 ms; field of view, 182×240 mm; slice thickness, 3 mm; spacing, 1 mm; matrix 88×116) with identical slice locations to those of transverse T2WI, and b values of 50, 1,000, and 1,500 s/mm^2. The apparent diffusion coefficient (ADC) value was calculated using the workstation with b values of 50 and 1,000 s/mm^2, and an ADC map was also generated.

Histopathological work-up

Prostatectomy specimens were treated as described in a previous study (21). Whole-mount tissue slices were stained with hematoxylin-eosin (HE) to create histologic slides (HE procedure stains the nucleus and cytoplasm contrasting colors to differentiate the fine structures of cells and tissues). Histological slices were assessed by a pathologist with 13 years of experience in urogenital pathology who was ignorant to the biopsy and MRI results.

Tumor foci were delineated on microscopy slides. Residual tumors were concordantly identified by two senior genitourinary pathologists. Pathological analysis was conducted independently from imaging analysis, and the pathologists were blinded to the imaging data in informing their tissue analysis. All prostate specimens were completely submitted to identify residual lesions.

Each residual cancer was labeled artificially on the glass

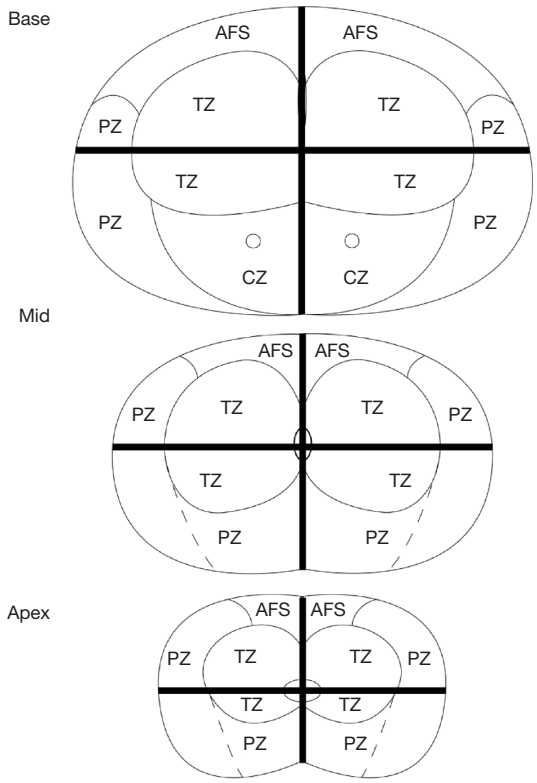

Figure 2 Diagrams of the prostate on data collection forms. AFS, anterior fibromuscular stroma; PZ, peripheral zone; TZ, transition zone; CZ, central zone.

slide of the whole-mount section and on the corresponding area of the 12-sector map, similar to a previous study (*Figure 2*) (12).

Image analysis

One radiologist with 5 years of experience in MRI organized the reading sessions on an Advantage Workstation 4.7 (GE HealthCare, Chicago, IL, USA), serving as the study coordinator, to obtain a final set of independent readings from the other three pelvic radiologists with different levels of experience in prostate imaging diagnosis (15, 8, and 3 years, respectively). Before the study, each reader had used PI-RR to interpret at least 100 post-ADT prostate MRI scans performed from 2013 to 2016.

MRI of the prostate was divided into 12 sectors, as described in the study by Kowa *et al.* (12). The three readers were aware of the inclusion criteria but ignorant of the history of patients. The PI-RR score was assessed based on sectors rather than on lesions because many lesions became scattered and small after ADT, making it difficult to confirm the number of lesions. Each sector was assessed and scored according to the PI-RR standard, and the content of score 5 was changed slightly (*Figure 3*) as follows: (I) no abnormality on high b value DWI (b value =1,500 s/mm^2), ADC map, and DCE image; (II) diffuse moderate hyperintensity on high b value DWI and/or diffuse moderate hypointensity on the ADC map, with DCE showing negative or slight enhancement; (III) focal marked hyperintensity on high b value DWI or focal marked hypointensity on the ADC map but not on both, with DCE showing negative or with slight enhancement; (IV) focal marked hyperintensity on high b value DWI and marked hypointensity on the ADC map but not in the same sectors of the primary tumor confirmed by biopsy or remarkable enhancement on DCE; and (V) focal marked hyperintensity on high b value DWI and marked hypointensity on the ADC map in the same sectors of the primary tumor confirmed by biopsy.

Upgrading from PI-RR 4 to 5 was appropriate if (I) the locations for diffusion restriction and enhancement matched and (II) if the radiologists judged the lesion in the same biopsy sector as the primary tumor (since the radiologists were blind to biopsy results).

Finally, any sector with imaging findings identified by readers (PI-RR >2) was scored using PI-RR and recorded on the sector map.

The study coordinator and pathologist reviewed sector maps to match histological residual lesions with the imaging findings. If a finding lay in the same sector or at least half of the lesion lay in the same sector, it was considered to be a match.

Statistical analysis

All of the statistical analyses were performed with dedicated software (Stata Statistical version 12; StataCorp, College Station, TX, USA). The percentage of sectors with residual cancer was calculated for each score, and Spearman correlation analysis was performed to assess the relationship between percentage and PI-RR score. The per-sector sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) with the detection of the sector with residual cancer were calculated at PI-RR category 3 and 4 thresholds. To determine sensitivity, the cancer detection rate (CDR) per sector was calculated as the ratio of the number of sectors with suspicious MRI findings ultimately confirmed to be PCa to the total number of sectors with residual lesions found on histology; this was

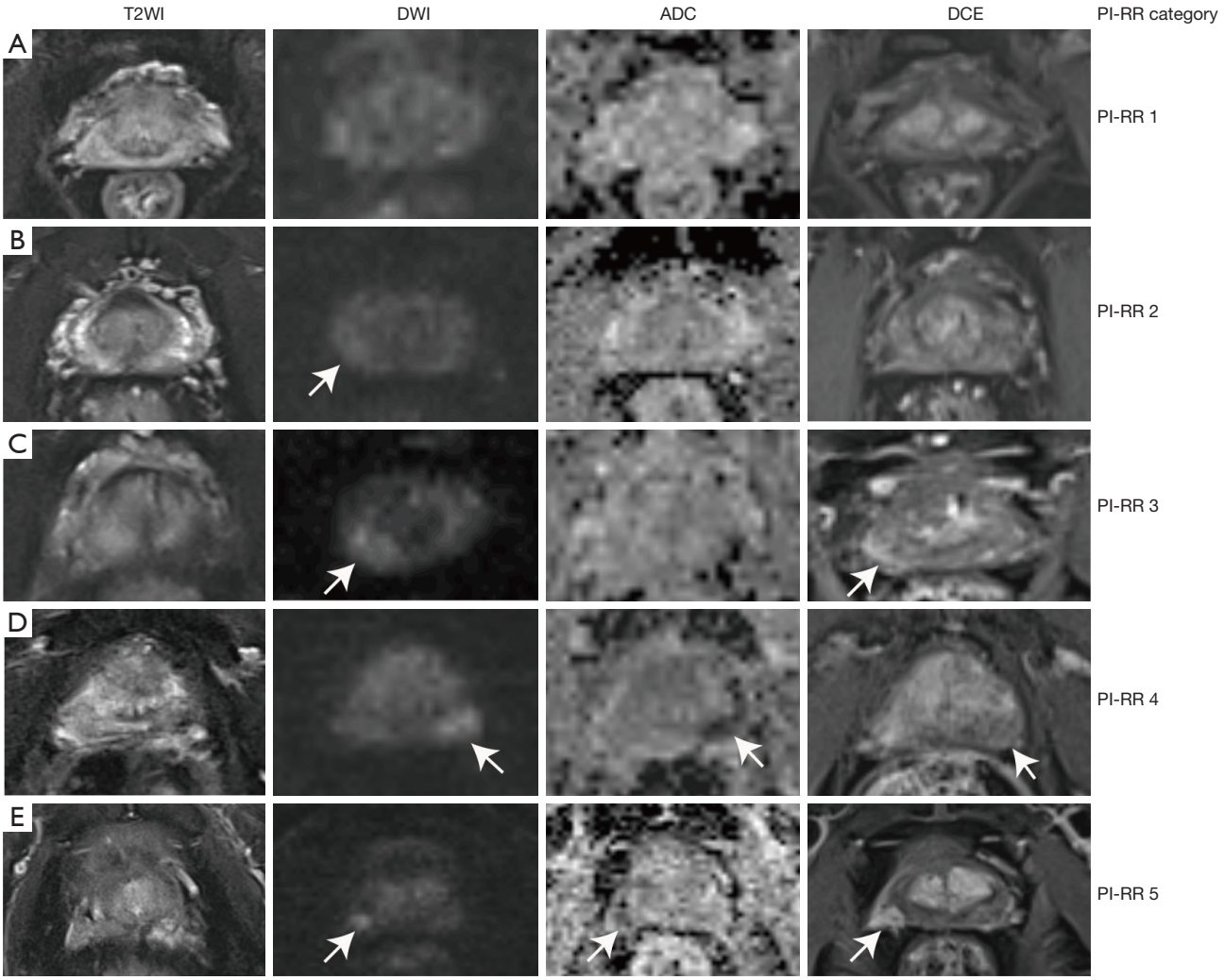

Figure 3 Examples of MRI after ADT. (A) Images from a 64-year-old man after ADT (PSA =0.01 ng/mL). MRI showed no abnormality on high b value DWI, ADC maps, or DCE imaging; therefore, all sectors were scored as PI-RR 1. (B) Images from a 64-year-old man (PSA =0.01 ng/mL). MRI showed diffuse and slight hyperintensity at DWI (arrow) on the right PZ with no suspicious focus of local relapse; hence, the sector was scored as PI-RR 2. (C) Images from a 69-year-old man (PSA =0.02 ng/mL). MRI showed focal hyperintensity on DWI (arrow) on the right PZ with slight enhancement (arrow) but no hypointensity on the ADC map; hence, the case was scored as PI-RR 3. (D) Images from a 66-year-old man (PSA =2.24 ng/mL). MRI revealed marked hyperintensity on the posterior part of the left PZ of the prostate on high b value DWI (arrow), with corresponding marked hypointensity on the ADC map and slight enhancement on DCE imaging (arrows); hence, the case was scored as PI-RR 4 (blind to biopsy results). (E) Images from an 85-year-old man (PSA =0.64 ng/mL). MRI showed remarkable hyperintensity on high b value DWI (arrow), hypointensity on the ADC map (arrow), and noticeable enhancement on DCE imaging (arrow) on the right PZ that matched the primary tumor sector confirmed by biopsy; hence, the sector was scored as PI-RR 5. T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; ADC, apparent diffusion coefficient; DCE, dynamic contrast-enhanced; PI-RR, Prostate Imaging for Recurrence Reporting; MRI, magnetic resonance imaging; ADT, androgen deprivation therapy; PSA, prostate-specific antigen; PZ, peripheral zone.

Table 1 Clinical characteristics of patients (n=119)

Characteristic	Values
Age (years), median [IQR]	68 [47–87]
PSA before ADT (ng/mL), median [IQR]	78.50 [4.63–1,410]
PSA after ADT (ng/mL), median [IQR]	0.09 [0.01–12.22]
Duration of ADT (months), median [IQR]	4.50 [3–51]
Gleason score, n (%)	
≤7	28 (23.53)
≥8	91 (76.47)
M stage, n (%)	
0	82 (68.91)
1	37 (31.09)

IQR, interquartile range; PSA, prostate-specific antigen; ADT, androgen deprivation therapy.

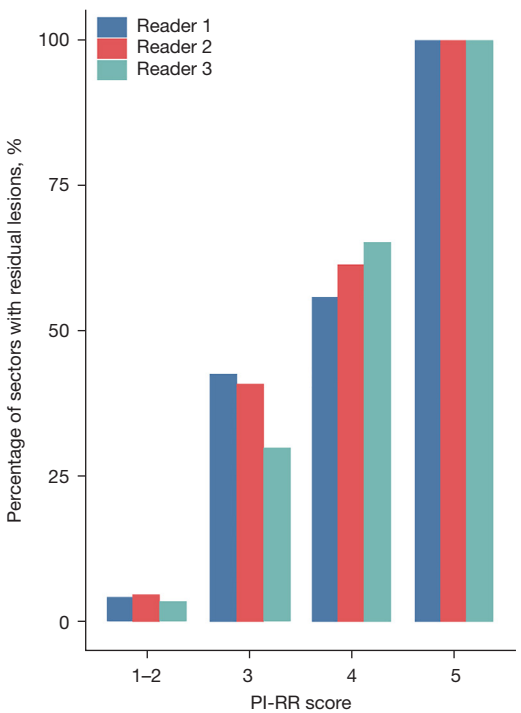

Figure 4 Percentage of sectors with residual lesions detected by readers. PI-RR, Prostate Imaging for Recurrence Reporting.

then then compared across the three readers with the chi-squared test. A suspicious sector was sectors with MRI findings of a category higher than 3 or 4.

Cohen's kappa statistic (κ) was used to assess the interreader agreement in attributing PI-RR categories of sectors with residual lesions, with MRI findings being classified as the presence or absence of residual lesions (using category 3 and 4 thresholds, respectively). κ was calculated using the analytical method in the case of dichotomous variables or bootstrapping when two readers were involved or the variables had two levels; a weighted κ was used if necessary. The 95% confidence interval (CI) was calculated with the bootstrap method. The reference values included slight agreement (0.01–0.20), fair agreement (0.21–0.40), moderate agreement (0.41–0.60), substantial agreement (0.61–0.80), and almost perfect agreement (0.81–0.99).

Results

Among the patients deemed eligible for inclusion, 80 with residual disease confirmed by pathology and 39 with pathological complete response (CR) were enrolled. All patients underwent a drastic decline in PSA levels, with the median PSA level dropping from 78.50 ng/mL [interquartile range (IQR), 4.63–1,410.00 ng/mL] to 0.09 ng/mL (IQR, 0.01–12.22 ng/mL) (P<0.05) after neoadjuvant treatment. The clinical characteristics of the patients are summarized in *Table 1*. The percentage of pathologically positive sectors increased as PI-RR score increased for all readers, and a significant correlation was observed between percentage and score (all P values <0.05) (*Figure 4*).

Diagnostic efficacy of the PI-RR at cutoffs of 3 and 4

The results of MRI readings are illustrated in *Tables 2,3*, in which the sensitivity, specificity, PPV, and NPV at cutoffs of 3 and 4 for readers 1–3 are also listed. With PI-RR 3 as the cutoff, among all readers, the sensitivity, specificity, PPV and NPV ranged from 74.2% to 83.7%, 86.4% to 92.7%, 51.3% to 64.3%, and 95.4% to 96.9%, respectively, and at a cutoff of 4, they ranged from 47.4% to 56.5%, 97.9% to 98.6%, 82.5% to 85.3%, and 91.6% to 92.9%, respectively. The CDR of readers with less experience was relatively higher, but there was no significantly difference among the CDR of three readers at the cutoff of score 3 or 4 (P>0.05).

Interreader agreement

We observed moderate agreement in assigning PI-RR classes and in evaluating of presence of residual cancer with a threshold of PI-RR ≥3 or PI-RR ≥4 among all readers (*Table 4* and *Figures 5,6*).

The agreement between more experienced readers

Table 2 Distribution of the PI-RR categories assigned by reader 1, reader 2, and reader 3 (15, 8, and 3 years of experience, respectively) and associated residual cancer detection performance

Lesions	Distribution		
	Reader 1	Reader 2	Reader 3
PI-RR assignment			
PI-RR 1	24	22	20
PI-RR 2	25	32	14
PI-RR 3	52	56	57
PI-RR 4	24	27	47
PI-RR 5	84	72	71
Sensitivity			
At cutoff 3	76.6% (160/209)	74.2% (155/209)	83.7% (175/209)
At cutoff 4	51.7% (108/209)	47.4 % (99/209)	56.5% (118/209)
Specificity			
At cutoff 3	92.7% (1,130/1,219)	92.0% (1,121/1,219)	86.4% (1,053/1,219)
At cutoff 4	98.4% (1,200/1,219)	98.6% (1,202/1,219)	97.9% (1,194/1,219)
PPV			
At cutoff 3	64.3% (160/249)	61.3% (155/253)	51.3% (175/341)
At cutoff 4	85.0% (108/127)	85.3% (99/116)	82.5% (118/143)
NPV			
At cutoff 3	95.8% (1,130/1,179)	95.4% (1,121/1,175)	96.9% (1,053/1,087)
At cutoff 4	92.2% (1,200/1,301)	91.6% (1,202/1,312)	92.9% (1,194/1,285)

PI-RR, Prostate Imaging for Recurrence Reporting; PPV, positive predictive value; NPV, negative predictive value.

Table 3 Confusion matrix of the PI-RR categories assigned by readers with variable experience

Actual	Predicted											
	Reader 1				Reader 2				Reader 3			
	Cutoff ≥3		Cutoff ≥4		Cutoff ≥3		Cutoff ≥4		Cutoff ≥3		Cutoff ≥4	
	P	N	P	N	P	N	P	N	P	N	P	N
P	160	49	108	101	155	54	99	110	175	34	118	91
N	89	1,130	19	1,200	98	1,121	17	1,202	166	1,053	25	1,194

PI-RR, Prostate Imaging for Recurrence Reporting; P, positive; N, negative.

(reader 1 *vs.* 2) was moderate in assigning PI-RR classes and substantial in evaluating the presence of cancer under both the PI-RR ≥3 and PI-RR ≥4 thresholds. The more experienced readers agreed with each other to a greater extent than when they were each compared with the less experienced reader (fair to moderate) (*Table 4*).

Discussion

In this study, the likelihood of existent residual cancer in sectors increased as the PI-RR score rose, and at the cutoff of a PI-RR score of 3, the experienced readers provided moderate sensitivity, specificity, PPV, and NPV, while at a

Table 4 Overall agreement and pairwise agreement between readers for all sectors with residual cancer and presence of residual cancer with a cutoff of PI-RR ≥3 or ≥4

Lesions	Overall agreement	Reader 1 *vs.* 2	Reader 2 *vs.* 3	Reader 1 *vs.* 3
PI-RR score	0.410 (0.371–0.450)	0.502 (0.428–0.575)	0.344 (0.272–0.415)	0.391 (0.318–0.463)
Presence of residual cancer				
At cutoff of PI-RR ≥3	0.495 (0.387–0.603)	0.616 (0.510–0.722)	0.419 (0.326–0.511)	0.445 (0.381–0.509)
At cutoff of PI-RR ≥4	0.521 (0.462–0.580)	0.694 (0.571–0.818)	0.449 (0.368–0.531)	0.423 (0.356–0.491)

κ values are reported together with 95% confidence interval in brackets. PI-RR, Prostate Imaging for Recurrence Reporting.

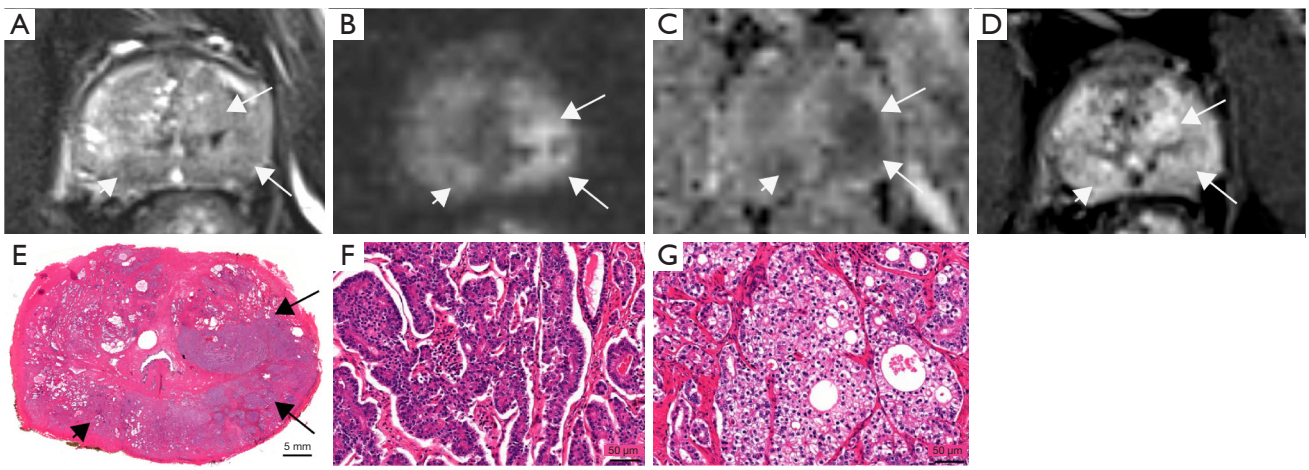

Figure 5 MRI scans in a 59-year-old man with residual cancer after ADT, in whom the initial PSA level was 41.22 ng/mL and was reduced to 10.04 ng/mL after 4-month therapy. (A) T2-weighted imaging showed a diffuse, low signal in the right PZ (short arrow), left PZ, and left TZ (long arrows). (B-D) DWI showed a remarkably high signal intensity and a low ADC value on the left TZ and PZ (long arrows), a slightly high signal on the right PZ (short arrows), remarkable enhancement on the left TZ, and slight enhancement on bilateral PZ. The three readers gave a PI-RR score of 5, 4, and 3 for the left TZ, left PZ, and right PZ, respectively (blind to the primary tumor site). (E) The residual cancer could be observed on the left TZ and PZ (long arrows) and right PZ (short arrow), stained with HE. (F) The left TZ showed residual cancer with significant malignant features (HE staining, ×5). (G) The lesions in right PZ and part of the left PZ showed slightly malignant characteristics after ADT, such as a clear cell pattern (HE staining, ×10). MRI, magnetic resonance imaging; ADT, androgen deprivation therapy; PSA, prostate-specific antigen; PZ, peripheral zone; TZ, transition zone; DWI, diffusion-weighted imaging; ADC, apparent diffusion coefficient; PI-RR, Prostate Imaging for Recurrence Reporting; HE, hematoxylin and eosin.

cutoff of 4, although the specificity, PPV and NPV were higher, the CDR/sensitivity was extremely low. For the less experienced reader, at a cutoff of 3 and 4, the CDR/sensitivity was relatively higher, but the PPV was extremely low.

There could be two explanations for this result. First, the T2WI signal shows different degrees of change due to the subsequent changes in hormonally treated prostate glands (22) and hampers zonal anatomy depiction. It is difficult to assess the existence of residual lesions using morphological features, and readers must rely on the change of DWI signal or ADC value. However, in normal tissue, ADC decreases after treatment due to ADT-induced reduction in the extravascular extracellular space, fibrosis, or apoptosis, which decreases the ADC value and DWI signal gap between the normal tissue and cancer or even mimics the residual lesions. Against such a complex background, identifying tumor lesions is extremely challenging, further resulting in overestimation for less experienced readers. Moreover, for the majority of PCas after ADT, the overall ADC value of the tumor would increase, and the volume would decrease, rendering the lesions less remarkable and further reducing the CDR.

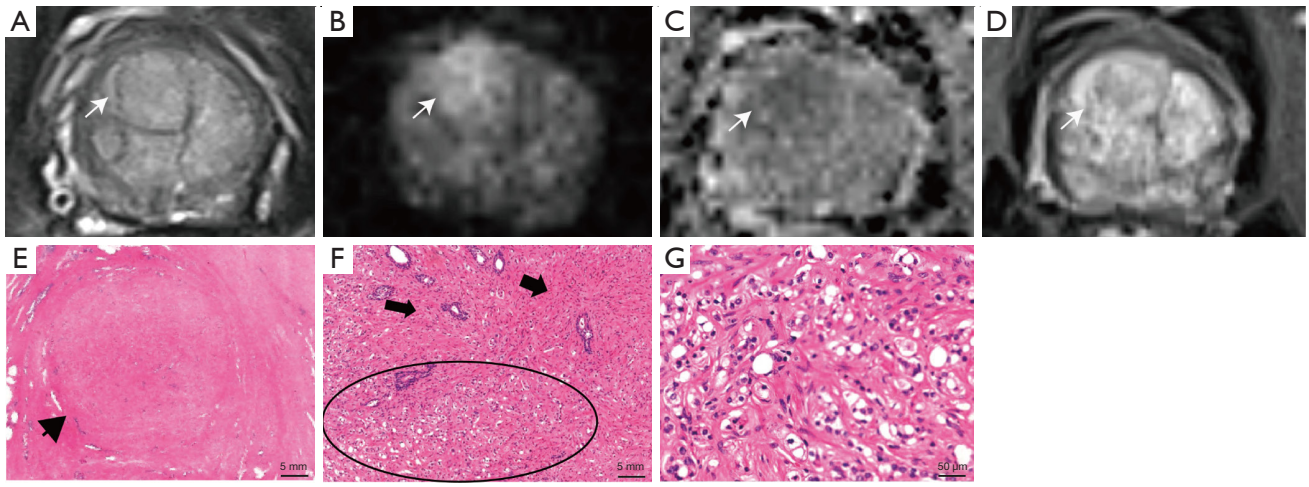

Figure 6 MRI scans in a 71-year-old man with residual cancer after ADT (PSA =0.61 ng/mL). There was disagreement among readers in the PI-RR scoring. (A) T2WI showed an elliptic lesion with a low signal (arrow) in the right anterior portion of the central gland. (B-D) The lesion displayed high signal intensity on DWI, a low ADC value, and moderate enhancement compared with the background (arrows). Reader 1 and 3 scored the images as PI-RR 4, but reader 2 gave a PI-RR of 2. The difference was because reader 1 and 3 judged the score mainly based on the DWI signal, ADC value, and DCE according to the PI-RR guide, while the reader 2 took account of the clear margin of the lesion on T2WI and considered it to be benign prostatic hyperplasia. (E) The lesion appeared as a nodule with a clear margin on the histopathological slice (arrow), stained with HE. (F) Histopathological imaging (HE staining, ×10) showed stromal hyperplasia (arrows) mixed with scattered residual cancer (circle). (G) On histopathological imaging (HE staining, ×10) the residual cancer showed malignant characteristics after ADT, such as nucleolus-poor clear cell pattern with clear cytoplasm and small hyperchromatic nuclei. MRI, magnetic resonance imaging; ADT, androgen deprivation therapy; PSA, prostate-specific antigen; PI-RR, Prostate Imaging for Recurrence Reporting; T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; ADC, apparent diffusion coefficient; DCE, dynamic contrast-enhanced; HE, hematoxylin and eosin.

In this study, readers with moderate to high experience showed moderate agreement in assigning PI-RR categories and moderate to substantial agreement in evaluating residual cancer with a cutoff of category 3 or 4. However, readers with low experience showed poor agreement with the other readers. Because of the structure and signal changes of normal tissues and lesions, it is difficult to establish a stable, clear morphological feature standard, such as the PI-RADS, for readers in terms of border, shape, or size (23). Moreover, cell apoptosis and atrophy of the gland occur at the same time during ADT, and histological change varies in the initial months of ADT, making the change in DWI signal and ADC value more uncertain and rendering it difficult to devise an optional cutoff for the signal or ADC value. On account of the lack of a clear standard, discrepancies between readers are practically unavoidable.

Our results are similar to those reported by Gold *et al.* (24), who reported that 13% of residual PCas post-ADT were negative and 75% weakly positive on DWI, indicating that a less strict standard would be suitable for the detection of residual diseases. The DWI detection of recurrence in the prostate after RT revealed similar results. Among studies (8,10,12) using restricted diffusion on both DWI and ADC maps as the standard of recurrence, similar to the score of 4 in our study, one study reported a sensitivity on a per sector basis (six sectors for every prostate) of 37.9% to 52.9% and a PPV of 66.7% to 76.7% (10), while another study (eight sectors for every prostate) reported a sensitivity of 68% and a PPV of 75% (8). Among studies using Likert scores, one reported a sensitivity ranging from 47% to 59% and a specificity from 82% to 93% per sector at a cutoff of likelihood ≥3, while at a cutoff of likelihood ≥4, the sensitivity ranged from 22% to 55% and the specificity from 87% to 95% (9). Overall, these studies mentioned above and our own indicate a need for a less strict standard for the detection of lesions after therapy and suggest that a PI-RR ≥3 should lead to biopsy. For PI-RR 4–5 findings, whose PPV is high and is associated with a greater likelihood of relapse, salvage therapy without biopsy could be indicated, whereas targeted biopsy could be used

to confirm relapse in clinically low-risk cases (2).

Several limitations to our study should be mentioned. First, although the prostate after ADT is similar to that after RT in its imaging and pathology characteristics, the pathological changes are more complex than are those after ADT. In addition to the influence being similar to that of ADT, the change in RT has a radiation-induced effect, which varies according to different doses and durations. Differences also exist between the two therapies in the application of the PI-RR; for instance, category 4 after RT might indicate a new lesion, but in this study, it mainly indicated a lesion missed by biopsy. Second, our findings might be difficult to generalize across different organizations because readers from the same institution tend to show similar approaches and attitudes to image interpretation (25). Prospective, multicenter studies with larger samples are required in the future. Third, we did not evaluate the reproducibility of ADC and DWI images or perform lesion size assessment.

Conclusions

When readers used the PI-RR system, we observed moderate agreement in assigning the categories and evaluating the spectrum of residual lesions found on whole-mount histology after ADT, and we found acceptable agreement among experienced radiologists. Although the interreader agreement and cancer diagnosis were influenced by readers' experience, the PI-RR system could effectively detect residual cancer on post-ADT MRI and might provide a potential reference for the care of PCa after RT.

Acknowledgments

Funding: This work was supported by the National Cancer Center of China (No. NCC201909B03) and the Clinical Science and Research Fund of Shanghai Municipal Health Commission (No. 2020040270).

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-23-1643/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-23-1643/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and solved. This retrospective single-center study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the Ethics Committee of Fudan University Shanghai Cancer Center. Informed consent was obtained from all individual participants.

References

1. Gupta RT, Mehta KA, Turkbey B, Verma S. PI-RADS: Past, present, and future. J Magn Reson Imaging 2020;52:33-53.
2. Panebianco V, Villeirs G, Weinreb JC, Turkbey BI, Margolis DJ, Richenberg J, Schoots IG, Moore CM, Futterer J, Macura KJ, Oto A, Bittencourt LK, Haider MA, Salomon G, Tempany CM, Padhani AR, Barentsz JO. Prostate Magnetic Resonance Imaging for Local Recurrence Reporting (PI-RR): International Consensus -based Guidelines on Multiparametric Magnetic Resonance Imaging for Prostate Cancer Recurrence after Radiation Therapy and Radical Prostatectomy. Eur Urol Oncol 2021;4:868-76.
3. Pecoraro M, Turkbey B, Purysko AS, Girometti R, Giannarini G, Villeirs G, Roberto M, Catalano C, Padhani AR, Barentsz JO, Panebianco V. Diagnostic Accuracy and Observer Agreement of the MRI Prostate Imaging for Recurrence Reporting Assessment Score. Radiology 2022;304:342-50.
4. Asuncion A, Walker PM, Bertaut A, Blanc J, Labarre M, Martin E, Bardet F, Cassin J, Cormier L, Crehange G, Loffroy R, Cochet A. Prediction of prostate cancer recurrence after radiation therapy using multiparametric magnetic resonance imaging and spectroscopy: assessment

of prognostic factors on pretreatment imaging. Quant Imaging Med Surg 2022;12:5309-25.
5. Franco PN, Frade-Santos S, García-Baizán A, Paredes-Velázquez L, Aymerich M, Sironi S, Otero-García MM. An MRI assessment of prostate cancer local recurrence using the PI-RR system: diagnostic accuracy, inter-observer reliability among readers with variable experience, and correlation with PSA values. Eur Radiol 2024;34:1790-803.
6. Light A, Kanthabalan A, Otieno M, Pavlou M, Omar R, Adeleke S, et al. The Role of Multiparametric MRI and MRI-targeted Biopsy in the Diagnosis of Radiorecurrent Prostate Cancer: An Analysis from the FORECAST Trial. Eur Urol 2024;85:35-46.
7. Shah TT, Kanthabalan A, Otieno M, Pavlou M, Omar R, Adeleke S, et al. Magnetic Resonance Imaging and Targeted Biopsies Compared to Transperineal Mapping Biopsies Before Focal Ablation in Localised and Metastatic Recurrent Prostate Cancer After Radiotherapy. Eur Urol 2022;81:598-605.
8. Tamada T, Sone T, Jo Y, Hiratsuka J, Higaki A, Higashi H, Ito K. Locally recurrent prostate cancer after high-dose-rate brachytherapy: the value of diffusion-weighted imaging, dynamic contrast-enhanced MRI, and T2-weighted imaging in localizing tumors. AJR Am J Roentgenol 2011;197:408-14.
9. Luzurier A, Jouve De Guibert PH, Allera A, Feldman SF, Conort P, Simon JM, Mozer P, Compérat E, Boudghene F, Servois V, Lucidarme O, Granger B, Renard-Penna R. Dynamic contrast-enhanced imaging in localizing local recurrence of prostate cancer after radiotherapy: Limited added value for readers of varying level of experience. J Magn Reson Imaging 2018;48:1012-23.
10. Donati OF, Jung SI, Vargas HA, Gultekin DH, Zheng J, Moskowitz CS, Hricak H, Zelefsky MJ, Akin O. Multiparametric prostate MR imaging with T2-weighted, diffusion-weighted, and dynamic contrast-enhanced sequences: are all pulse sequences necessary to detect locally recurrent prostate cancer after radiation therapy? Radiology 2013;268:440-50.
11. Liao XL, Wei JB, Li YQ, Zhong JH, Liao CC, Wei CY. Functional Magnetic Resonance Imaging in the Diagnosis of Locally Recurrent Prostate Cancer: Are All Pulse Sequences Helpful? Korean J Radiol 2018;19:1110-8.
12. Kowa JY, Soneji N, Sohaib SA, Mayer E, Hazell S, Butterfield N, Shur J, Ap Dafydd D. Detection and staging of radio-recurrent prostate cancer using multiparametric MRI. Br J Radiol 2021;94:20201423.
13. Zattoni F, Kawashima A, Morlacco A, Davis BJ, Nehra AK, Mynderse LA, Froemming AT, Jeffrey Karnes R. Detection of recurrent prostate cancer after primary radiation therapy: An evaluation of the role of multiparametric 3T magnetic resonance imaging with endorectal coil. Pract Radiat Oncol 2017;7:42-9.
14. Menges D, Yebyo HG, Sivec-Muniz S, Haile SR, Barbier MC, Tomonaga Y, Schwenkglenks M, Puhan MA. Treatments for Metastatic Hormone-sensitive Prostate Cancer: Systematic Review, Network Meta-analysis, and Benefit-harm assessment. Eur Urol Oncol 2022;5:605-16.
15. Magi-Galluzzi C, Sanderson H, Epstein JI. Atypia in nonneoplastic prostate glands after radiotherapy for prostate cancer: duration of atypia and relation to type of radiotherapy. Am J Surg Pathol 2003;27:206-12.
16. Bostwick DG, Meiers I. Diagnosis of prostatic carcinoma after therapy. Arch Pathol Lab Med 2007;131:360-71.
17. Koopman AGMM, Jenniskens SFM, Fütterer JJ. Magnetic Resonance Imaging Assessment After Therapy in Prostate Cancer. Top Magn Reson Imaging 2020;29:47-58.
18. Potretzke TA, Froemming AT, Gupta RT. Post-treatment prostate MRI. Abdom Radiol (NY) 2020;45:2184-97.
19. Patel P, Mathew MS, Trilisky I, Oto A. Multiparametric MR Imaging of the Prostate after Treatment of Prostate Cancer. Radiographics 2018;38:437-49.
20. Vargas HA, Wassberg C, Akin O, Hricak H. MR imaging of treated prostate cancer. Radiology 2012;262:26-42.
21. Chen ZZ, Gu WJ, Zhou BN, Liu W, Gan HL, Zhang Y, Zhou LP, Liu XH. Radiomics based on biparametric MRI for the detection of significant residual prostate cancer after androgen deprivation therapy: using whole-mount histopathology as reference standard. Asian J Androl 2023;25:86-92.
22. Chen M, Hricak H, Kalbhen CL, Kurhanewicz J, Vigneron DB, Weiss JM, Carroll PR. Hormonal ablation of prostatic cancer: effects on prostate morphology, tumor detection, and staging by endorectal coil MR imaging. AJR Am J Roentgenol 1996;166:1157-63.
23. Rudolph MM, Baur ADJ, Haas M, Cash H, Miller K, Mahjoub S, Hartenstein A, Kaufmann D, Rotzinger R, Lee CH, Asbach P, Hamm B, Penzkofer T. Validation of the PI-RADS language: predictive values of PI-RADS lexicon descriptors for detection of prostate cancer. Eur Radiol 2020;30:4262-71.
24. Gold SA, VanderWeele DJ, Harmon S, Bloom JB, Karzai F, Hale GR, Marhamati S, Rayn KN, Mehralivand S, Merino MJ, Gulley JL, Bilusic M, Madan RA, Choyke PL, Turkbey B, Dahut W, Pinto PA. mpMRI preoperative

staging in men treated with antiandrogen and androgen deprivation therapy before robotic prostatectomy. Urol Oncol 2019;37:352.e25-30.
25. Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, Westphalen AC, Babb JS, Margolis DJ. Interobserver Reproducibility of the PI-RADS Version 2 Lexicon: A Multicenter Study of Six Experienced Prostate Radiologists. Radiology 2016;280:793-804.

Cite this article as: Chen Z, Zhou B, Liu W, Gan H, Chen R, Yang L, Zhou L, Liu X. Diagnostic efficacy and interobserver agreement among readers with variable experience of the Prostate Imaging for Recurrence Reporting system with whole-mount histology after androgen deprivation therapy as a reference. Quant Imaging Med Surg 2024;14(4):3006-3017. doi: 10.21037/qims-23-1643

