European Journal of Radiology 121 (2019) 108704

Contents lists available at ScienceDirect

European Journal of Radiology

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

Comparison of PI-RADS version 2 and PI-RADS version 2.1 for the detection of transition zone prostate cancer

Tsutomu Tamada[a,*], Ayumu Kido[a], Mitsuru Takeuchi[b], Akira Yamamoto[a], Yoshiyuki Miyaji[c], Naoki Kanomata[d], Teruki Sone[a]

[a] Department of Radiology, Kawasaki Medical School, Kurashiki, Japan
[b] Department of Radiology, Radiolonet Tokai, Nagoya, Japan
[c] Department of Urology, Kawasaki Medical School, Kurashiki, Japan
[d] Department of pathology, Kawasaki Medical School, Kurashiki, Japan

ARTICLE INFO

Keywords:
Prostate cancer
MR imaging
Multiparametric MRI
Transition zone
Prostate Imaging Reporting and Data System Version 2
Prostate Imaging Reporting and Data System Version 2.1

ABSTRACT

Purpose: To compare the diagnostic performance of PI-RADS v2 and v2.1 for detecting transition zone prostate cancer (TZPC) on multiparametric prostate MRI (mpMRI).

Method: Fifty-eight patients with elevated PSA levels underwent mpMRI at 3 T including T2-weighted imaging (T2WI) and diffusion-weighted imaging (DWI), and subsequent MRI–transrectal ultrasonography fusion-guided prostate-targeted biopsy (MRGB). The standard of reference was MRGB-derived histopathology. Two readers independently assessed each TZ lesion, assigning a score of 1–5 for T2WI, a score of 1–5 for DWI, and the overall PI-RADS assessment category according to PI-RADS v2 and v2.1. The diagnostic performance of the two methods was compared in terms of inter-reader agreement, diagnostic sensitivity, diagnostic specificity, and area under the ROC curve (AUC).

Results: Of the 58 patients, 26 were diagnosed with PC (GS = 3 + 3, n = 9; GS = 3 + 4, n = 9; GS = 3 + 5, n = 1; GS = 4 + 3, n = 4; GS = 4 + 4, n = 3) and 32 with benign lesions. Regarding inter-reader agreement of overall PI-RADS assessment category, the kappa value was 0.580 for v2 and 0.645 for v2.1. For both readers, there was no difference in diagnostic sensitivity between the versions ($p \geq 0.500$). For reader 1, the diagnostic specificity was higher for v2.1 ($p = 0.002$), and was similar for reader 2 ($p = 1.000$). For both readers, AUC tended to be higher for v2.1 than for v2, but the difference was not significant (0.786 vs. 0.847 for reader 1, $p = 0.052$; and 0.808 vs. 0.858 for reader 2, $p = 0.197$).

Conclusions: These results suggest that compared with PI-RADS v2, PI-RADS v2.1 could be preferable for evaluating TZ lesions.

1. Introduction

In 2018, ~164,690 new cases of prostate cancer (PC) were diagnosed in the United States and ~78,400 in Japan, and PC is the second and the sixth leading cause of cancer death in men in these countries [1,2].

Although the majority of PC tumors occur in the peripheral zone (PZ), ~30% arise in the transition zone (TZ) [3]. Prostate multiparametric magnetic resonance imaging (mpMRI) clearly depicts the zonal anatomy of the prostate, including PZ and TZ, and is currently the most accurate diagnostic imaging tool for detecting tumors in either zone. [4,5]. In 2015, the American College of Radiology, European Radiology of Uroradiology, and AdMeTech Foundation jointly released the Prostate Imaging and Reporting Data System Version 2 (PI-RADS v2), which was standardized for assessment of the probability (PI-RADS assessment category) of clinically significant PC (csPC) using prostate

Abbreviations: PI-RADS v2, Prostate Imaging Reporting and Data System Version 2; PI-RADS v2.1, Prostate Imaging Reporting and Data System Version 2.1; TZPC, transition zone prostate cancer; mpMRI, multiparametric prostate MRI; T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; MRGB, MRI–transrectal ultrasonography fusion-guided prostate-targeted biopsy; PC, prostate cancer; PZ, peripheral zone; TZ, transition zone; csPC, clinically significant PC; BPH, benign prostatic hyperplasia; PSA, prostate specific antigen; TRUS, transrectal ultrasonography; GS, Gleason score; 3D, three-dimensional; TSE, turbo spin echo; TR, repetition time; TE, echo time; FOV, field of view; MPG, motion-probing gradient; ADC, apparent diffusion coefficient; PPV, positive predictive value; NPV, negative predictive value: AUC the area under the curve; ROC, receiver operating characteristic; csTZPCs, clinically significant TZPCs

* Corresponding author at: Department of Radiology, Kawasaki Medical School, 577 Matsushima, Kurashiki city, Okayama, 701-0192, Japan.
E-mail address: ttamada@med.kawasaki-m.ac.jp (T. Tamada).

https://doi.org/10.1016/j.ejrad.2019.108704
Received 22 July 2019; Received in revised form 3 October 2019; Accepted 12 October 2019

mpMRI [6]. However, PI-RADS v2 has several limitations, including sub-optimal inter-reader reproducibility; relatively high false negative rates for the low PI-RADS assessment category; lower detection rates for TZ tumors than for PZ tumors; and problems with decision rules, including ambiguous assessment criteria for TZ on T2-weighted imaging (T2WI) [7–17]. For example, for TZ lesions, all lesions except normal TZ (which is rare in elderly patients with suspected PC) (T2WI score 1), typical glandular benign prostatic hyperplasia (BPH) nodule (T2WI score 2), and highly suspicious TZ lesion (T2WI scores 4 and 5) are classified as T2WI score 3 [6]; however, a large number of atypical lesions are also classified as T2WI score 3, including BPH (mainly stromal BPH) to TZPC, which is difficult to differentiate on T2WI [11,18,19].

To address limitations in PI-RADS v2, the PI-RADS Steering Committee developed an updated version (PI-RADS v2.1) in 2019 [20]. One of the major modifications in the updated version is the diagnostic criteria for tumors in the TZ on T2WI (mainly for low T2WI score). It is anticipated that the clinical use of PI-RADS v2.1 will improve inter-reader variability and further simplify the PI-RADS assessment of prostate MRI [20]. However, to our knowledge, no previous studies have evaluated inter-reader reproducibility and tumor detection ability in TZPC using PI-RADS v2.1. The purpose of this study was to compare the diagnostic performance between PI-RADS v2 and PI-RADS v2.1 for TZPC.

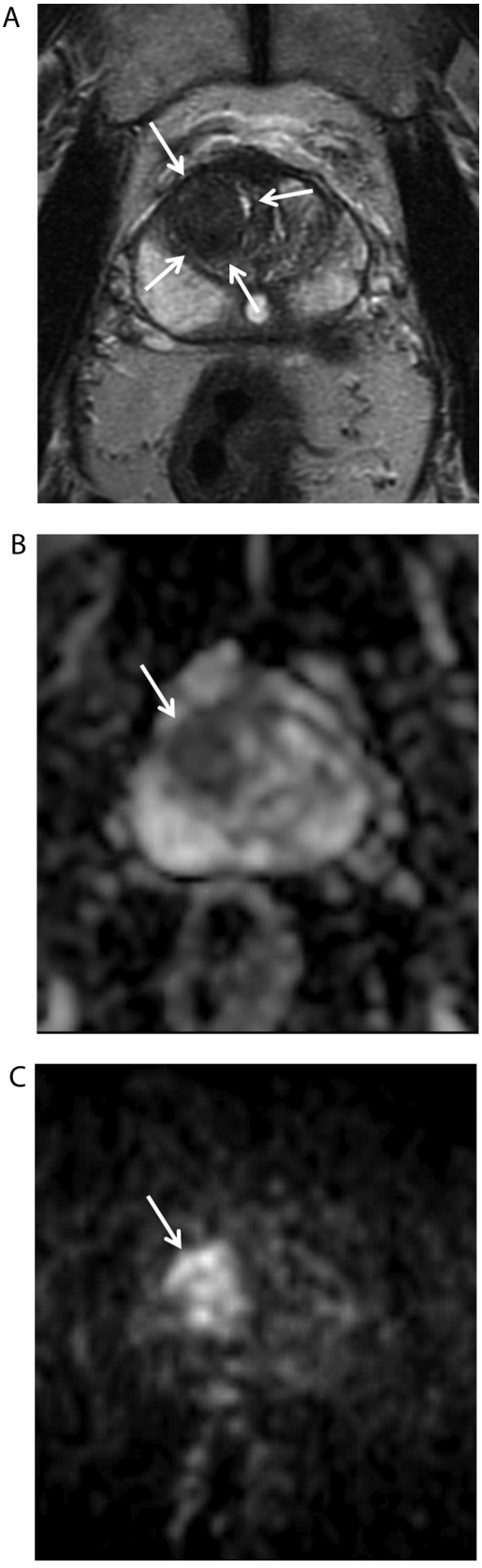
2. Materials and methods

2.1. Study population

Our Institutional Review Board approved this retrospective study and waived the need for informed consent. Between August 2018 and March 2019, 97 consecutive patients with elevated prostate specific antigen (PSA) levels underwent 3 T prostate mpMRI and subsequent MRI–transrectal ultrasonography (TRUS) fusion-guided prostate-targeted biopsy (MRGB) for lesions suspicious for PC on mpMRI. Thirty-nine patients were removed from the study based on the following exclusion criteria: suspected PZ lesions on mpMRI (n = 37); insignificant TZPC with Gleason score (GS) = 3 + 3 and tumor size < 0.5 mL (maximum tumor diameter < 8 mm on mpMRI) (n = 1); and a period of 6 months or more between MRI and MRGB (n = 1). Thus, 58 patients (age range, 45–87 years; mean age, 69.7 years) with suspected TZPC were included in the study. The mean PSA level at the time of the initial MRI was 8.07 ± 4.98 ng/mL (PSA range, 4.07–33.69 ng/mL; median PSA level, 6.54 ng/mL). The mean interval between the MRI examination and MRGB was 36 ± 19 days (range, 2–104 days; median, 37 days). No patient had undergone any therapy for PC at the time of the MRI examination.

2.2. MRGB

All MRGB procedures were performed using the UroStation system (Koelis; Grenoble, France) with elastic image fusion, real-time three-dimensional (3D) tracking technology, and a computer workstation (Koelis) for segmentation of the prostate and the lesion under local sacral and TRUS-guided periprostatic plexus anesthesia [21]. Prior to the biopsy, a radiologist performed segmentation of the whole prostate and MRI-defined lesions from 3D mpMR image data on the UroStation workstation. The mpMRI 3D volume data and the real-time TRUS image were then elastically fused on the screen. Immediately after the biopsy core of the target lesion displayed on the fusion image was obtained by a biopsy needle under TRUS guidance, additional real-time 3D TRUS images were obtained to determine the accuracy of needle deployment within the target lesion. MRGB was performed for suspicious lesions according to PI-RADS v2 and conventional overall multiparametric MRI assessment before this study [6,22]. At least two cores were obtained for each MRI-targeted lesion. Additional cores were obtained at the



(caption on next page)

Fig. 1. A 71-year-old man with suspected prostate cancer (prostate-specific antigen level, 5.00 ng/mL) in the right transition zone. MRI-transrectal ultrasonography fusion-guided prostate-targeted biopsy revealed a lesion with features consistent with prostate cancer (Gleason score = 3 + 3; diameter, 18.1 mm). a: T2-weighted image shows an area of heterogeneous signal intensity that does not qualify as score 2, 4, or 5, located in the middle region of the transition zone in PI-RADS v2 (reader 1), and a heterogeneous encapsulated nodule in the middle region of the transition zone in PI-RADS v2 (reader 2); in PI-RADS v2.1, both readers described the nodule as mostly encapsulated (arrows). b: Apparent diffusion coefficient (ADC) map shows marked hypointensity according to both readers (arrow). c: Diffusion-weighted image shows marked hyperintensity according to both readers (arrow). The lesion was assigned a T2-weighted imaging score of 3, a DWI/ADC map score of 5, and overall category of 3 in PI-RADS v2 by reader 1; and a T2-weighted imaging score of 2, a DWI/ADC map score of 5, and overall category of 2 in PI-RADS v2 by reader 2. For both readers, the lesion was assigned a T2-weighted imaging score of 2, a DWI/ADC map score of 5, and an overall category of 3 according to the PI-RADS v2.1 decision rules. Thus, compared with PI-RADS v2, PI-RADS v2.1 improved inter-reader agreement and also the diagnostic performance for transition zone prostate cancer detection in reader 2.

Table 1
Inter-reader Agreement of PI-RADS Assessment Category between Two Readers for PI-RADS v2 and PI-RADS v2.1.

	Reader 2						
	PI-RADS v2 assessment category	1	2	3	4	5	Total
Reader 1	1	0	0	0	0	0	0
	2	0	8	0	0	0	8
	3	0	10	22	3	0	35
	4	0	0	3	8	0	11
	5	0	0	0	0	4	4
	Total	0	18	25	11	4	58
	Reader 2						
	PI-RADS v2.1 assessment category	1	2	3	4	5	Total
Reader 1	1	5	1	0	0	0	6
	2	2	7	5	0	0	14
	3	2	0	19	3	0	24
	4	0	0	2	8	0	10
	5	0	0	0	0	4	4
	Total	9	8	26	11	4	58

discretion of the operator, based on lesion size, lesion location, and the confidence in targeting accuracy.

2.3. Histopathologic examination

The MRGB specimens underwent hematoxylin–eosin staining. A uropathologist with 23 years of experience and who was blinded to the MRI findings recorded the presence or absence of PC, tumor size index (total lesion core length/total core length), and GS of the tumor, for each specimen. The GS was evaluated according to the 2014 International Society of Urological Pathology Modified Gleason Grading System [23]. A lesion was considered clinically significant PC when GS was ≥ 7 and tumor diameter ≥ 5 mm, or with GS = 3 + 3 and tumor size ≥ 0.5 mL (tumor diameter ≥ 8 mm). Tumor size was calculated from the mpMR images, most commonly T2WI.

2.4. MR imaging technique

The MRI examinations were performed using a 3 T scanner with a 32-channel phased-array coil (Ingenia 3.0 T CX Quasar Dual; Philips Medical Systems, Best, The Netherlands). All examinations were performed with patients in the fasting condition, and all patients received intramuscular buscopan® (SANOFI, Tokyo, Japan) or glucagon to reduce intestinal peristalsis.

The following pulse sequences were used: axial turbo spin echo (TSE) T2WI [repetition time (TR)/echo time (TE) 7257/95 ms; slice thickness 3 mm; no intersection gap; field of view (FOV) 200 × 200 mm, matrix 352 × 277, in-plane resolution 0.57 × 0.72 mm^2, parallel imaging factor 1.4], coronal TSE T2WI [TR/TE 5730/95 ms; slice thickness 4 mm; no intersection gap; FOV 200 × 200 mm, matrix 272 × 187, in-plane resolution 0.74 × 1.07 mm^2, parallel imaging factor 1.8], and axial single-shot spin echo echo-planar diffusion-weighted imaging (DWI) [TR/TE 6000/70 ms; slice thickness 3 mm; no intersection gap; FOV 300 × 300 mm, matrix 112 × 112, in-plane resolution 2.68 × 2.68 mm^2, parallel imaging factor 2; b values 0, 1000, and 2000s/mm^2). DWI was acquired with motion-probing gradient (MPG) pulses applied sequentially along the three orthogonal orientations following acquisition at 3 b values of 0–2000 s/mm^2. Apparent diffusion coefficient (ADC) maps were reconstructed by calculating the ADC in each pixel of each slice, and ADC values were calculated for a pair of b values (0 and 2000s/mm^2) by a mono-exponential fitting. Axial T1WI, 3D T2WI, and dynamic contrast-enhanced MR imaging were also performed, but not assessed in the present study. The 3D T2WI was used for MRI–TRUS fusion-guided prostate-targeted biopsy, since it improve the accuracy of elastic fusion between 3D MR data and 3D TRUS data.

2.5. Image analysis

Prior to assessment by the radiologists, a study coordinator prepared a PowerPoint file showing the mpMR images of the TZ lesion targeted for MRGB, for each patient. Two fellowship-trained abdominal radiologists, with 7 and 12 years of experience in prostate MRI independently assessed the mpMRI of each TZ lesion using PI-RADS v2

Table 2
Comparison of Diagnostic Performance for Transition Zone Prostate Cancer Detection between PI-RADS v2 and PI-RADS v2.1.

	Reader 1			Reader 2		
	PI-RADS		P value	PI-RADS		P value
	v 2	v 2.1		v 2	v 2.1	
Sensitivity	100 (26/26)	92.3 (24/26)	0.500	88.5 (23/26)	96.2 (25/26)	0.625
Specificity	25.0 (8/32)	56.3 (18/32)	0.002	46.9 (15/32)	50.0 (16/32)	1.000
Positive predictive value	52.0 (26/50)	63.2 (24/38)	NA	57.5 (23/40)	61.0 (25/41)	NA
Negative predictive value	100 (8/8)	90.0 (18/20)	NA	83.3 (15/18)	94.1 (16/17)	NA
Accuracy	58.6 (34/58)	72.4 (42/58)	0.039	65.5 (38/58)	70.7 (41/58)	0.508
AUC	0.786	0.847	0.053	0.808	0.858	0.197

Note- AUC = area under the receiver operating characteristic curve; NA = not applicable.
Data excluding AUC are percentages, with values used to calculate percentages in parentheses.
Positive predictive value = number of true positive / number of true positive + number of false positive.
Negative predictive value = number of true negative / number of true negative + number of false negative.

Table 3
Discrepant patients of Transition Zone Prostate Cancer Detection Results between PI-RADS v2 and PI-RADS v2.1 for Each Reader.

Outcome	PI-RADS v2 (v2) result	PI-RADS v2.1 (v2.1) result	Reader 1	Reader 2	Total number of patients
Patients with combination of incorrect in v2 and correct in v2.1	FP	TN	10	3	16
	FN	TP	0	3	
Patients with combination of correct in v2 and incorrect in v2.1	TP	FN	2	1	5
	TN	FP	0	2	

Note- TP = true positive; TN = true negative; FN = false negative; FP = false positive.
Data are number of patients.
A PI-RADS assessment category ≥3 of both PI-RADS versions was defined as positive for transition prostate cancer detection.

and PI-RADS v2.1. The mpMRI of each patient was presented on a dedicated workstation (Synapse EX, Fujifilm Corporation, Japan) and viewed with reference to the PowerPoint file. Prior to the present study, the radiologists had 3 years of experience in image interpretation using PI-RADS v2. They were aware of the patient's age and PSA level, but were blinded to the histopathology results. The radiologists independently assigned each tumor a score of 1–5 for T2WI, a score of 1–5 for DWI, and an overall PI-RADS assessment category according to PI-RADS v2 and PI-RADS v2.1 [6,20]. In PI-RADS v2.1, assessment of T2WI was performed using axial and coronal images because PI-RADS v2.1 requires the features of the shape and margin of TZ to be assessed in at least two planes on T2WI [20]. Detailed information regarding the assessment categories and algorithms of both PI-RADS versions is listed in Supplemental Tables 1–4.

2.6. Statistical analysis

Statistical analysis was performed at the 5% significance level using SPSS for Windows v. 22.0 software (SPSS, Chicago, IL) and JMP v. 11.0.0 software (SAS, Cary, NC). Inter-reader agreement of the PI-RADS assessment category was evaluated for PI-RADS v2 and PI-RADS v2.1 using the kappa statistic. Kappa values < 0.20 indicated poor agreement, 0.21–0.40 indicated fair agreement, 0.41–0.60 indicated moderate agreement, 0.61–0.80 indicated good agreement, and ≥0.81 indicated excellent agreement. Diagnostic sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) for TZPC detection were calculated for PI-RADS v2 and PI-RADS v2.1. In calculating the diagnostic performance, data were dichotomized according to a predefined cut-off value: for both PI-RADS versions, a PI-RADS assessment category ≥3 was defined as positive. Regarding intra-reader agreement, differences in sensitivity, specificity, and accuracy between the two PI-RADS versions for each radiologist were tested using the McNemar test. The area under the curve (AUC) of receiver operating characteristic (ROC) analysis was used to evaluate the diagnostic performance for TZPC detection in both PI-RADS versions. Comparison of the AUC between the two PI-RADS versions was performed for each radiologist, using the Delong test.

3. Results

Of the 58 patients, 26 were diagnosed with csPC and 32 with benign lesions. The 26 TZPCs comprised GS = 3 + 3, n = 9; GS = 3 + 4, n = 9; GS = 3 + 5, n = 1; GS = 4 + 3, n = 4; and GS = 4 + 4, n = 3. The mean diameter of the TZPC was 12.3 mm ± 3.8 mm (range, 7.9–21.0 mm).

Regarding inter-reader agreement of the PI-RADS assessment category between the two readers, the kappa value was 0.580 for PI-RADS v2 (indicating moderate agreement) and 0.645 for PI-RADS v2.1 (indicating good agreement) (Fig. 1). In PI-RADS v2, there was disagreement in 10 patients (assessment category 3 in reader 1 and assessment category 2 in reader 2) (Table 1). In PI-RADS v2.1, there was disagreement in 7 patients (assessment category 3 in reader 1 and assessment category 1 in reader 2, n = 2; assessment category 2 in reader 1 and assessment category 3 in reader 2, n = 5) (Table 1). The proportions of patients in the five assessment categories were closer between the two readers in PI-RADS v2.1 than in PI-RADS v2 (Table 1). For reader 1, 12 of 35 patients of assessment category 3 in PI-RADS v2 were reclassified to category 2 in PI-RADS v2.1. For reader 2, 14 of 18 patients of assessment category 2 in PI-RADS v2 were reclassified to category 1 (n = 9) or to category 3 (n = 5) in PI-RADS v2.1; and 4 of 25 patients of assessment category 3 in PI-RADS v2 were reclassified to category 2 in PI-RADS v2.1.

In comparing the diagnostic performance for TZPC detection between PI-RADS v2 and PI-RADS v2.1, there was no difference in diagnostic sensitivity for either reader (100% (26/26 cases) vs. 92.3% (24/26 cases), p = 0.500; and 88.5% (23/26 cases) vs. 96.2% (25/26 cases), p = 0.625, respectively) (Table 2). The diagnostic specificity and accuracy were higher for v2.1 for reader 1 (25.0% (8/32 cases) vs. 56.3% (18/32 cases), p = 0.002; and 58.6% (34/58 cases) vs. 72.4% (42/58 cases), p = 0.039, respectively), and similarly for reader 2 (46.9% (15/32 cases) vs. 50.0% (16/32 cases), p = 1.000; and 65.5% (38/58 cases) vs. 70.7% (41/58 cases), p = 0.508, respectively) (Table 2). AUC tended to be higher in v2.1 than in v2 for both readers, but the difference was not significant (0.786 vs. 0.847 for reader 1, p = 0.053; and 0.808 vs. 0.858 for reader 2, p = 0.197) (Table 2). In discrepant patients of TZPC detection results between PI-RADS v2 and PI-RADS v2.1 for both readers, 16/116 patients (13.8%) with 13 false positive results (all category 3) in PI-RADS v2 for both readers (10 in reader 1 and 3 in reader 2) and 3 false negative results (all category 2) for reader 2 in PI-RADS v2 were identified as a true negative result (all category 2) and a true positive result (all category 3: score 2 in T2WI and score 4 in DWI (new diagnostic algorithm in PI-RADS v2.1 (Online material 4))) in PI-RADS v2.1, respectively (Table 3) (Figs. 1 and 2). Conversely, 5/116 patients (4.3%) with 3 true positive results (all category 3) in PI-RADS v2 for both readers (2 in reader 1 and 1 in reader 2) and 2 true negative results (all category 2) in PI-RADS v2 for reader 2 were identified as a false negative result (all category 2) and a false positive result (all category 3: score 2 in T2WI and score 4 in DWI (Online material 4)) in PI-RADS v2.1, respectively (Table 3). In addition, 28/116 patients (24.1%) for both readers showed a false positive result in both PI-RADS versions (14 patients in reader 1 and 14 patients in reader 2).

4. Discussion

We compared the inter-reader reproducibility and diagnostic performance for TZPC detection between PI-RADS v2 and PI-RADS v2.1, using MRGB as the reference standard. Our study comprised 58 patients, including 32 benign TZ lesions and 26 clinically significant TZPCs (csTZPCs) with a wide range of tumor aggressiveness (GS = 3 + 3 (low risk), n = 9; GS = 3 + 4 (intermediate risk), n = 9; and GS ≥ 4 + 3 (high risk), n = 8). Therefore, we consider that the study population was suitable for evaluating the diagnostic performance for TZPC detection.

PI-RADS v2.1 had better inter-reader reproducibility than did PI-RADS v2 (good agreement and moderate agreement, respectively) (not proven statistically). There was a disagreement between the two readers regarding various combinations of findings. Among them, a

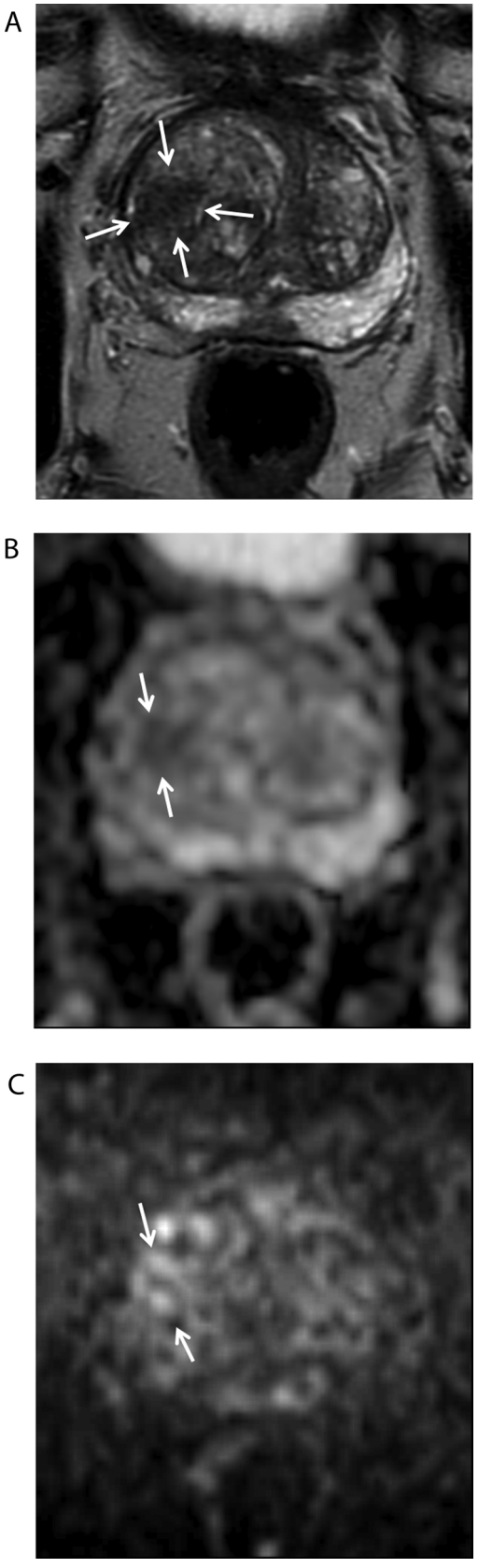


Fig. 2. A 78-year-old man with suspected prostate cancer (prostate-specific antigen level, 7.18 ng/mL) in the right transition zone. MRI-transrectal ultrasonography fusion-guided prostate-targeted biopsy confirmed a diagnosis of benign prostatic condition with hyperplasia and atrophic glands. a: T2-weighted image shows an area of heterogeneous hypointensity with obscured margins, including some that do not qualify as score 2, 4, or 5, in the middle lateral region of the transition zone in PI-RADS v2, and a mostly encapsulated nodule in PI-RADS v2.1 (arrows). b: Apparent diffusion coefficient (ADC) map shows indistinct mildly hypointensity (arrows). c: Diffusion-weighted image shows indistinct mild hyperintensity (arrows). According to the PI-RADS v2 decision rules, the lesion was assigned a T2-weighted imaging score of 3, a DWI/ADC map score of 2, and the overall PI-RADS category was 3 for both readers. For both readers, the lesion was assigned a T2-weighted imaging score of 2, a DWI/ADC map score of 2 (including some that do not qualify as 3, 4, or 5), and an overall category of 2 according to the PI-RADS v2.1 decision rules. Thus, compared with PI-RADS v2, PI-RADS v2.1 improved the diagnostic performance for transition zone prostate cancer detection in both readers.

combination of a positive result for TZPC and a negative results for TZPC (e.g., a positive result (assessment category of 3) in reader 1 and a negative result (assessment category of 2) in reader 2) is problematic in the management of patients with suspected PC including indication of MRGB [24]. It should be noted that the number of combinations of such different results between the two readers was lower in PI-RADS v2.1 (n = 7) than in PI-RADS v2 (n = 10). Furthermore, the proportions of patients in the five assessment categories were closer between the two readers in PI-RADS v2.1 than in PI-RADS v2 (Table 1). The reason for this might be that many patients with a PI-RADS assessment category of 2 or 3 in PI-RADS v2 for both readers were properly reclassified in PI-RADS v2.1 because of the more detailed diagnostic criteria for scores 1 and 2 of T2WI in PI-RADS v2.1, and because DWI diagnostic criteria are mandatory for a PI-RADS assessment category of 3 [20].

In comparing the diagnostic performance for TZPC detection between the two PI-RADS versions, compared with PI-RADS v2, the use of PI-RADS v2.1 significantly increased the specificity in reader 1, and increased the sensitivity in reader 2 (not significant). In particular, the reason why the specificity of PI-RADS v2.1 improved in reader 2 compared to reader 1 is that the specificity of PI-RADS v2 in reader 1 is lower than that of reader 2 (25.0% vs. 46.9%). This is probably due to the low inter-reader reproducibility of PI-RADS v2. Although the accuracy and AUC were higher in PI-RADS v2.1 than in PI-RADS v2 for both readers, these differences were significant only for accuracy in reader 1. PI-RADS v2.1 correctly diagnosed 16 lesions that were false positive or false negative for both readers in PI-RADS v2; in contrast, PI-RADS v2.1 misdiagnosed 5 patients that both readers had diagnosed correctly in PI-RADS v2. These findings indicate that in the present cohort, PI-RADS v2.1 has improved tumor detection ability for csTZPC compared with PI-RADS v2. The detailed diagnostic criteria in PI-RADS v2.1 of low T2WI score (including internal and perilesional properties) and the incorporation of DWI criteria in determination of the PI-RADS assessment category for an atypical lesion with T2WI score 2 may allow further improvement of the differential diagnosis between glandular BPH (typical BPH), stromal BPH, and TZPC using mpMRI [11,18–20]. However, the relationships between lesions of assessment categories 1–3 in PI-RADS v2.1 and the histopathological changes should be evaluated to support the improved diagnostic performance of PI-RADS v2.1. It is of note that PI-RADS v2.1 still showed a high false positive rate for csTZPC detection, similar to that with PI-RADS v2 [25]. Therefore, further modification of assessment categories 3–5 in PI-RADS is warranted to improve detection of csTZPC.

There are several limitations in this study. First, this was a retrospective, single-center study with a relatively small number of patients and statistically underpowered. Therefore, the present results may need further validation in prospective multi-center studies with a larger number of patients. Second, the reference standard was MRGB, which allows a broader spectrum of patients to be included than the patient cohort that undergoes radical prostatectomy, including patients without prostate cancer [26]. In addition, previous studies have demonstrated that the combination of MRGB and systematic biopsy

reliably detects 97% of csPC lesions when radical prostatectomy is used as the reference standard, and showed 95% overall negative predictive value for ruling out csPC [27,28]. Third, we assessed the diagnostic performance of PI-RADS v2.1 only for TZPC detection although in the clinical setting, radiologists also evaluate PZ lesions, which show higher morbidity in PC. However, in this study, we focused on TZPC detection using PI-RADS v2.1, for the reasons that PI-RADS v2 has several clinical problems with lower cancer detection rates for diagnosis of TZ tumors compared with PZ tumors, and because the decision rules include ambiguous assessment criteria for T2WI of the TZ. Finally, we did not investigate the relationship between pathological assessments and detailed MRI findings for the low PI-RADS v2.1 assessment category.

5. Conclusions

In conclusion, the present findings suggest that PI-RADS v2.1 improves inter-reader reproducibility (not proven statistically) and may contribute to increased diagnostic performance in terms of diagnostic specificity, accuracy, and AUC (not significant). There was a significant or marginally significant difference in TZPC detection with PI-RADS v2.1 compared with PI-RADS v2 in reader 1, which was unrelated to the reader's experience with PI-RADS. Thus, we consider that compared with PI-RADS v2, PI-RADS v2.1 is a more suitable method for detecting csPC in the TZ prior to prostate biopsy. Revisions to PI-RADS have steadily achieved standardization of qualitative assessment using mpMRI for csPC in the TZ.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors have no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2019.108704.

References

[1] American Cancer Society, Cancer facts and Figures, (2018) (Accessed 4 June 2019), https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2018/cancer-facts-and-figures-2018.pdf.
[2] Center for Cancer Control and Information Services, Projected Cancer Statistics, National Cancer Center Japan, 2018 (Accessed 4 June 2019), https://ganjoho.jp/en/public/statistics/short_pred.html.
[3] J.E. McNeal, E.A. Redwine, F.S. Freiha, T.A. Stamey, Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread, Am. J. Surg. Pathol. 12 (12) (1998) 897–906.
[4] C.M. Hoeks, J.O. Barentsz, T. Hambrock, et al., Prostate cancer: multiparametric MR imaging for detection, localization, and staging, Radiology 261 (1) (2011) 46–66.
[5] H.U. Ahmed, A. El-Shater Bosaily, L.C. Brown, et al., Diagnostic accuracy of multiparametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study, Lancet 389 (10078) (2017) 815–822.
[6] J.C. Weinreb, J.O. Barentsz, P.L. Choyke, et al., PI-RADS prostate imaging - reporting and data system: 2015, Version 2, Eur. Urol. 69 (1) (2016) 16–40.
[7] A.B. Rosenkrantz, L.A. Ginocchio, D. Cornfeld, et al., Interobserver reproducibility of the PIRADS Version 2 lexicon: a multicenter study of six experienced prostate radiologists, Radiology 280 (3) (2016) 793–804.
[8] S. Borofsky, A.K. George, S. Gaur, et al., What are we missing? False-negative cancers at multiparametric MR imaging of the prostate, Radiology 286 (1) (2018) 186–195.
[9] T. Auer, M. Edlinger, J. Bektic, et al., Performance of PI-RADS version 1 versus version 2 regarding the relation with histopathological results, World J. Urol. 35 (5) (2017) 687–693.
[10] A. Kido, T. Tamada, N. Kanomata, et al., Multidimensional analysis of clinicopathological characteristics of false-negative clinically significant prostate cancers on multiparametric MRI of the prostate in Japanese men, J. Radiol. 37 (2) (2019) 154–164.
[11] A.L. Chesnais, E. Niaf, F. Bratan, et al., Differentiation of transitional zone prostate cancer from benign hyperplasia nodules: evaluation of discriminant criteria at multiparametric MRI, Clin. Radiol. 68 (6) (2013) e323–330.
[12] B.G. Muller, J.H. Shih, S. Sankineni, et al., Prostate cancer: interobserver agreement and accuracy with the revised Prostate Imaging Reporting and Data System at multiparametric MR imaging, Radiology 277 (3) (2015) 741–750.
[13] A.B. Rosenkrantz, J.S. Babb, S.S. Taneja, J.M. Ream, Proposed adjustments to PI-RADS version 2 decision rules: impact on prostate cancer detection, Radiology 283 (1) (2017) 119–129.
[14] A.B. Rosenkrantz, A. Oto, B. Turkbey, A.C. Westphalen, Prostate imaging reporting and data system (PI-RADS), version 2: a critical look, AJR Am. J. Roentgenol. 206 (6) (2016) 1179–1183.
[15] S. Polanec, T.H. Helbich, H. Bickel, et al., Head-to-head comparison of PI-RADS v2 and PI-RADS v1, Eur. J. Radiol. 85 (2016) 1125–1131.
[16] M. Benndorf, F. Hahn, M. Krönig, et al., Diagnostic performance and reproducibility of T2w based and diffusion weighted imaging (DWI) based PI-RADSv2 lexicon descriptors for prostate MRI, Eur. J. Radiol. 93 (2017) 9–15.
[17] M. Benndorf, L. Waibel, M. Krönig, et al., Peripheral zone lesions of intermediary risk in multiparametric prostate MRI: frequency and validation of the PI-RADSv2 risk stratification algorithm based on focal contrast enhancement, Eur. J. Radiol. 99 (2018) 62–67.
[18] T. Tamada, T. Sone, Y. Jo, et al., Apparent diffusion coefficient values in peripheral and transition zones of the prostate: comparison between normal and malignant prostatic tissues and correlation with histologic grade, J. Magn. Reson. Imaging 28 (3) (2008) 720–726.
[19] A. Oto, A. Kayhan, Y. Jiang, et al., Prostate cancer: differentiation of central gland cancer from benign prostatic hyperplasia by using diffusion-weighted and dynamic contrast-enhanced MR imaging, Radiology 257 (3) (2010) 715–723.
[20] B. Turkbey, A.B. Rosenkrantz, M.A. Haider, et al., Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. 2019, Eur. Urol. (2019), https://doi.org/10.1016/j.eururo.2019.02.033.
[21] S. Verma, P.L. Choyke, S.C. Eberhardt, et al., The current state of MR imaging-targeted biopsy techniques for detection of prostate cancer, Radiology 285 (2) (2017) 343–356.
[22] T. Tamada, T. Sone, H. Higashi, et al., Prostate cancer detection in patients with total serum prostate-specific antigen levels of 4-10 ng/mL: diagnostic efficacy of diffusion-weighted imaging, dynamic contrast-enhanced MRI, and T2-weighted imaging, AJR Am. J. Roentgenol. 197 (3) (2011) 664–670.
[23] J.I. Epstein, L. Egevad, M.B. Amin, B. Delahunt, J.R. Srigley, P.A. Humphrey, Grading Committee. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system, Am. J. Surg. Pathol. 40 (2) (2016) 244–252.
[24] A.B. Rosenkrantz, S. Verma, P. Choyke, et al., Prostate magnetic resonance imaging and magnetic resonance imaging targeted biopsy in patients with a prior negative biopsy: a consensus statement by AUA and SAR, J. Urol. 196 (6) (2016) 1613–1618.
[25] F.V. Mertan, M.D. Greer, J.H. Shih, et al., Prospective evaluation of the prostate imaging reporting and data system version 2 for prostate cancer detection, J. Urol. 196 (3) (2016) 690–696.
[26] T. Tamada, J.M. Ream, A.M. Doshi, S.S. Taneja, A.B. Rosenkrantz, Reduced field-of-View diffusion-weighted magnetic resonance imaging of the prostate at 3 tesla: comparison with standard echo-planar imaging technique for image quality and tumor assessment, J. Comput. Assist. Tomogr. 41 (6) (2017) 949–956.
[27] J.P. Radtke, C. Schwab, M.B. Wolf, et al., Multiparametric magnetic resonance imaging (MRI) and MRI-transrectal ultrasound fusion biopsy for index tumor detection: correlation with radical prostatectomy specimen, Eur. Urol. 70 (5) (2016) 846–853.
[28] L. Boesen, N. Nørgaard, V. Løgager, H.S. Thomsen, Clinical outcome following low suspicion multiparametric prostate magnetic resonance imaging or benign magnetic resonance imaging guided biopsy to detect prostate cancer, J. Urol. 198 (2) (2017) 310–315.