Received:
11 June 2017

Revised:
4 August 2017

Accepted:
8 August 2017

<https://doi.org/10.1259/bjr.20170438>

Cite this article as:

Kwon M-R, Kim CK, Kim J-H. PI-RADS version 2: evaluation of diffusion-weighted imaging interpretation between $b = 1000$ and $b = 1500$ s mm⁻². *Br J Radiol* 2017; **90**: 20170438.

FULL PAPER

PI-RADS version 2: evaluation of diffusion-weighted imaging interpretation between $b = 1000$ and $b = 1500$ s mm⁻²

¹MI-RI KWON, MD, ^{1,2}CHAN KYO KIM, MD and ¹JAE-HUN KIM, PhD

¹Department of Radiology and Center for Imaging Science, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

²Department of Medical Device Management and Research, SAIHST, Sungkyunkwan University, Seoul, Korea

Address correspondence to: Dr Chan Kyo Kim
E-mail: chankyokim@skku.edu

Objective: To investigate the variability of diffusion-weighted imaging (DWI) interpretation of Prostate Imaging Reporting and Data System (PI-RADS) version 2 (v2) in evaluating prostate cancer (PCa).

Methods: 154 patients with PCa underwent multiparametric 3T MRI, followed by radical prostatectomy. DWI with different b values ($b = 0, 100, 1000$ and 1500 s mm⁻²) was obtained. Using the PI-RADS v2, two radiologists independently scored suspicious lesions in each patient and compared DWI of $b = 1000$ (DWI₁₀₀₀) with 1500 (DWI₁₅₀₀) s mm⁻².

Results: On DWI₁₀₀₀ and DWI₁₅₀₀, the intermethod and interobserver agreements of DWI scores were excellent in all patients ($\kappa \geq 0.873$). In each peripheral zone and transition zone DWI scores, both observers showed

excellent intermethod agreement between DWI₁₀₀₀ and DWI₁₅₀₀ ($\kappa \geq 0.897$), and interobserver agreement for DWI₁₀₀₀ and DWI₁₅₀₀ was good to excellent ($\kappa \geq 0.796$). For estimating clinically significant cancer, the area under receiver operating characteristics curves of DWI₁₀₀₀ and DWI₁₅₀₀ were 0.710 and 0.724 for observer 1 ($p = 0.11$), and 0.649 and 0.656 for observer 2 ($p = 0.12$), respectively.

Conclusion: The PI-RADS v2 scoring at 3T shows excellent agreement between DWI₁₀₀₀ and DWI₁₅₀₀ in evaluating PCa, with excellent inter-observer agreement.

Advance in knowledge: DWI using $b = 1000$ s mm⁻² instead of $b = 1500$ s mm⁻² reduces examination time or image distortion, with improved the signal-to-noise ratio.

INTRODUCTION

The recently published Prostate Imaging Reporting and Data System (PI-RADS) version 2 (v2) was developed to improve global standardization and reduce variation in image acquisition, interpretation, and reporting of multiparametric MRI (mpMRI).¹ The comprehensive report aims to simplify and standardize MRI interpretation using a straightforward framework that can be applied in daily practice. Several recent studies have reported moderate to good interobserver agreement in the peripheral zone (PZ) using the PI-RADS v2 scoring scales.^{2,3}

According to the PI-RADS v2 guidelines, the dominant MRI sequence in the PZ of the prostate is diffusion-weighted imaging (DWI) regardless of the assessment of other pulse sequences, except that positive dynamic contrast-enhanced MRI (DCE-MRI) score upgrades the final score from 3 to 4.¹ In the transition zone (TZ), the dominant MRI sequence is T_2 weighted imaging (T2WI) and a role of DWI is that score 5 upgrades the final score from 3 to 4. The guidelines

introduced and recommended to routinely use the “high b -values ≥ 1400 s mm⁻²” on DWI, if the signal-to-noise (SNR) is adequate. These recommendations are supported by the results of several recent studies that DWI using $b = 1500$ – 2500 s mm⁻² was optimal for the detection of prostate cancer (PCa) and superior to conventional DWI using $b = 1000$ s mm⁻² (DWI₁₀₀₀).^{4–11} Although such high b -value DWI confers greater benign tissue suppression and improves the detection of clinically significant cancer (CSC), it is limited because it reduces the SNR, increases the susceptibility artefacts and distorts the images due to increased eddy currents.¹² Furthermore, high b -value DWI ≥ 1400 s mm⁻² can be computed or synthesized by extrapolation using lower b -value data, which is acquired to develop the apparent diffusion coefficient (ADC) map; however this requires additional software and more validated results. Currently, numerous institutions that perform prostate mpMRI use conventional DWI ($b = 800$ – 1000 s mm⁻²) in daily practice. Therefore, it is questionable whether higher b -value DWI (>1000 s mm⁻²) is markedly different from

conventional DWI₁₀₀₀ in terms of the PI-RADS v2 scoring scale. The purpose of the present study was to investigate the inter-method and interobserver variabilities between DWI₁₀₀₀ and DWI with a b -value of 1500 s mm^{-2} (DWI₁₅₀₀) in terms of the PI-RADS v2 scoring.

Patients

The Institutional Review Board of our institution approved this retrospective study, and waived the requirement for informed consent. From August 2015 to July 2016, 188 patients with biopsy-proven PCa that received the radical prostatectomy were enrolled in our study. Of these 188 patients, 34 were excluded for the following reasons: outside MRI examinations ($n = 27$) or MRI examinations without DWI₁₅₀₀ ($n = 3$), preoperative androgen deprivation therapy ($n = 2$), and prostate sarcoma ($n = 2$). Accordingly, 154 patients (mean age, 64 years; range, 43–78 years) met the following inclusion criteria: (1) preoperative mpMRI at 3T including DWI₁₀₀₀ and DWI₁₅₀₀, (2) no previous treatment before the surgery and (3) surgically proven adenocarcinoma. The mean prostate-specific antigen (PSA) level was 6.1 ± 9.7 (standard deviation) ng/ml. The mean time interval between MRI examination and surgery was 33.4 ± 24.6 days.

Table 1. Clinical and pathologic characteristics ($n = 154$)

Characteristic	Data
Age (year)	64 (43–78)
Time interval between MRI and surgery (day)	33.4 (1–171)
PSA (ng ml ⁻¹)	6.1 (0.6–54.9)
Median Gleason score	7 (6–9)
6	13
7	113
8	9
9	19
Cancer volume (cm ³)	5.51 (0.01–81.8)
Clinical significant cancer (n)	146
Cancer location (n)	
PZ	106
TZ	48
Pathologic stage (n)	
2a	11
2b	1
2c	85
3a	38
3b	18
4	1

Note: PSA, prostate-specific antigen; PZ, peripheral zone; TZ, transition zone.

Unless otherwise indicated, data are means with the range in parentheses.

MR technique

All MRI was performed using a 3T MR system (Intera Achieva 3TX, Philips Medical System, Best, The Netherlands) equipped with a 6-channel SENSE coil (Philips Healthcare, Best, The Netherlands). All patients underwent T_1 weighted imaging, T2WI, DWI and DCE-MRI according to the guidelines of the European Society of Urogenital Radiology.¹³ T_2 weighted turbo spin-echo images were acquired in three orthogonal planes (transverse, sagittal, and coronal). The imaging parameters were as follows: repetition time/echo time, 3300–4000/100 ms; slice thickness, 3 mm; interslice gap, 1 mm; 568×341 matrix; field of view (FOV), 20 cm; number of signals acquired, 3; sensitivity encoding (SENSE) factor, 2; and acquisition time of each plane, 3 min 52–56 s. Transverse T_1 weighted turbo spin-echo images (slice thickness, 4 mm; and acquisition time, 3 min 12 s) were acquired to assess the lymph nodes and the pelvic bone. Transverse DCE-MRI was obtained using a 3D-fast field echo sequence (repetition time/echo time, 7.4/3.9 ms; flip angle, 5° and 15° [precontrast] and 25° [postcontrast]; matrix, 224×179 ; slice thickness, 4 mm; interslice gap, no; NSA, 1; FOV, 20 cm; 11 partitions on a 3D slab; and acquisition time, 7 min 10 s). A postcontrast series was performed immediately after a bolus injection of gadobutrol (Gadovist, Bayer Healthcare) at a dose of 0.1 mmol kg^{-1} body weight and a rate of $2\text{--}3 \text{ ml s}^{-1}$ using a power injector, followed by a 20 ml saline flush.

DWI was acquired in the transverse plane using the single-shot echo-planar imaging technique with parallel imaging and fat suppression (TR/TE, 5250/68 ms; slice thickness, 3 mm; interslice gap, 1 mm; matrix, 124×121 ; FOV, 20 cm; SENSE, 2; NSA, 3; and acquisition time, 5 min 4 s). Diffusion-encoding gradients were applied as 4 different b values (0, 100, 1000 and 1500 s mm^{-2}). The ADC maps were automatically developed on a pixel-by-pixel basis on both DWI₁₀₀₀ ($b = 100$ and 1000 s mm^{-2}) and DWI₁₅₀₀ ($b = 100$ and 1500 s mm^{-2}).

Image analysis

Two independent radiologists with 11 years (C.K.K.) and 1 year (M.K.) of experience in prostate MRI who were unaware of the clinical, surgical, and histological findings assessed the MR images.

For qualitative analysis, the prostate was divided into 18 regions: 12 in the PZ (medial and lateral regions of the base, mid-gland, and apex on the right and left side) and 6 in the TZ (right and left regions of the base, mid-gland and apex).

Instead of using the PI-RADS score in all regions of the prostate, we recorded the PI-RADS score per patient using an index lesion and a 5-point scale that was based on the PI-RADS v2 scoring scales. The index lesion was defined as a lesion identified on mpMRI with the highest PI-RADS assessment category and if the highest PI-RADS score was assigned to ≥ 2 lesions, they should be one that show extraprostatic extension or is larger. The likelihood of CSC according to the PI-RADS score criteria was determined as follows¹: very low (CSC is highly unlikely to be present; score 2, low (CSC is unlikely to be present); score 3, intermediate (the presence of CSC is equivocal); score 4, high

(CSC is likely to be present); and score 5, very high (CSC is highly likely to be present). The scores of DWI in both PZ and TZ were defined as the followings: score 1 (no abnormality on ADC and high b-value DWI), score 2 (indistinct hypointense on ADC), score 3 (focal mildly/moderately hypointense on ADC and isointense/mildly hyperintense on high b-value DWI), score 4 (focal markedly hypointense on ADC and markedly hyperintense on high b-value DWI with <1.5 cm in the greatest dimension), and score 5 (same as 4 but ≥1.5 cm in the greatest dimension or definite extraprostatic extension/invasive behaviour).

The PI-RADS v2 score for mpMRI including T2WI, DWI and DCE-MRI were reviewed by the two independent observers in two separate sessions. First, they assessed the PI-RADS score on mpMRI using the DWI₁₀₀₀. More than 2 weeks after the first assessment, they independently evaluated the PI-RADS score on mpMRI using the DWI₁₅₀₀. For TZ lesions, DCE-MRI was not assessed because the PI-RADS v2 assessment categories do not incorporate DCE-MRI in the TZ. If a lesion was located in both the PZ and TZ, the final PI-RADS score was recorded as the score of dominant area of the lesion. If the PZ was the dominant area of such a lesion, the final PI-RADS score was recorded as a PZ assessment. The greatest transverse diameter of the index lesion was also recorded in accordance with the strategy for lesion measurement of the PI-RADS v2.

Histopathological analysis

The findings in a whole-mounted step section of the resected prostate were used as the standard reference. All slides prepared from the tissue slices were reviewed by a pathologist with 12 years of experience in PCa who had no knowledge of the MRI findings; tumour size, volume, location, extracapsular extension, seminal vesicle invasion, and Gleason score were recorded.

Statistical analysis

To determine the DWI and final PI-RADS scores, the inter-method agreement between DWI₁₀₀₀ and DWI₁₅₀₀ and interobserver agreement for DWI₁₀₀₀ and DWI₁₅₀₀ were evaluated using κ statistics. A κ value of 0–0.20 was considered to indicate poor agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 good agreement, and 0.81–1.0 excellent agreement.¹⁴

The receiver operating characteristic curves were generated for estimating the presence of CSC using DWI₁₀₀₀ and DWI₁₅₀₀ in both observers. Diagnostic performance was then assessed by calculating the area under the curve (AUC). The sensitivity and specificity were evaluated for the presence of CSC regarding imaging data set and observers. The best threshold of sensitivity and specificity for estimating CSC using the PI-RADS v2 score was investigated; when the PI-RADS v2 score was ≥4, CSC was considered “present”.

Table 2. Distribution of DWI and final PI-RADS v2 scores on DWI₁₀₀₀ and DWI₁₅₀₀

		Observer 1				Observer 2			
DWI score		2	3	4	5	2	3	4	5
DWI ₁₀₀₀									
	All	39	35	42	38	28	43	37	46
	PZ	23	25	34	24	17	30	30	29
	TZ	16	10	8	14	11	13	7	17
DWI ₁₅₀₀									
	All	33	28	53	40	28	26	48	52
	PZ	20	17	43	26	17	18	38	33
	TZ	13	11	10	14	11	8	10	19
Final PI-RADS score									
Session 1									
	All	28	13	67	46	22	20	57	55
	PZ	21	8	53	24	17	8	52	29
	TZ	7	5	14	22	5	12	5	26
Session 2									
	All	25	12	69	48	22	19	54	59
	PZ	18	7	55	26	17	7	49	33
	TZ	7	5	14	22	5	12	5	26

Note: PZ, peripheral zone; TZ, transition zone.

Session 1 is the final PI-RADS score including DWI₁₀₀₀.

Session 2 is the final PI-RADS score including DWI₁₅₀₀.

The CSC was defined in cases of a Gleason score ≥ 7 , and/or volume $\geq 0.5 \text{ cm}^3$ and/or extracapsular extension.

Statistical analyses were performed using SPSS software (version 23.0; SPSS, Chicago, IL) and MedCalc (version 13.0; MedCalc Software, Mariakerke, Belgium). A p -value < 0.05 was considered statistically significant.

RESULTS

The patient and pathologic characteristics are summarized in Table 1. Among 154 patients, 146 (94.8%) had CSC and 8 (5.2%) lacked CSC. Among the 146 patients with CSC, 141 had a surgical Gleason score ≥ 7 , 135 had a tumour volume $\geq 0.5 \text{ cm}^3$, 38 had T3a cancer, 18 had T3b cancer and one had T4 cancer.

Regarding DWI and the final PI-RADS score, the distributions of the two observers are shown in Table 2. On the DWI1000 images,

Table 3. Intermethod and Interobserver agreement of DWI and final PI-RADS score on DWI₁₀₀₀ and DWI₁₅₀₀

Comparison	Location	κ value	95% CI
DWI ₁₀₀₀ and DWI ₁₅₀₀			
DWI score in observer 1	All	0.918	0.863–0.973
	PZ	0.890	0.805–0.975
	TZ	0.964	0.932–0.996
DWI score in observer 2	All	0.905	0.850–0.961
	PZ	0.897	0.825–0.969
	TZ	0.920	0.834–1.000
Final PI-RADS in observer 1	All	0.911	0.837–0.986
	PZ	0.864	0.751–0.977
	TZ	1.000	1.000–1.000
Final PI-RADS in observer 2	All	0.979	0.953–1.000
	PZ	0.967	0.926–1.000
	TZ	1.000	1.000–1.000
Observer 1 and 2			
DWI ₁₀₀₀ score	All	0.873	0.824–0.923
	PZ	0.890	0.840–0.941
	TZ	0.844	0.742–0.946
DWI ₁₅₀₀ score	All	0.870	0.813–0.927
	PZ	0.913	0.866–0.961
	TZ	0.796	0.667–0.925
Final PI-RADS on DWI ₁₀₀₀	All	0.771	0.667–0.874
	PZ	0.856	0.779–0.932
	TZ	0.600	0.347–0.853
Final PI-RADS on DWI ₁₅₀₀	All	0.803	0.703–0.903
	PZ	0.910	0.860–0.959
	TZ	0.600	0.347–0.853

Note: DWI, diffusion-weighted imaging; PI-RADS, Prostate Imaging Reporting and Data System; PZ, peripheral zone; TZ, transition zone.

observer 1 conferred the following scores: score 2—39 patients, score 3—5 patients, score 4—42 patients and score 5—38 patients; observer 2 provided the following scores: score 2—28 patients, score 3—30 patients, score 4—30 patients and score 5—37 patients. On the DWI1500 images, observer 1 conferred the following scores: score 2—33 patients, score 3—28 patients, score 4—53 patients and score 5—40 patients; observer 2 gave the following scores: score 2—28 patients, score 3—26 patients, score 4—48 patients and score 5—52 patients.

Table 3 presents the intermethod and interobserver agreements of the DWI and final PI-RADS scores determined on DWI₁₀₀₀ and DWI₁₅₀₀ images. Regarding DWI scores in all patients, the intermethod agreement between DWI₁₀₀₀ and DWI₁₅₀₀ was excellent ($\kappa = 0.918$ for observer 1 and 0.905 for observer 2; Figures 1 and 2). The interobserver agreement for DWI₁₀₀₀ and DWI₁₅₀₀ was excellent ($\kappa = 0.873$ for DWI₁₀₀₀ and 0.870 for DWI₁₅₀₀). For the subanalysis for each PZ and TZ DWI scores, the intermethod agreements between DWI₁₀₀₀ and DWI₁₅₀₀ images were excellent for both observer ($\kappa \geq 0.897$ in all cases), and interobserver agreements for DWI₁₀₀₀ and DWI₁₅₀₀ images were good to excellent ($\kappa \geq 0.796$ in all cases). The final PI-RADS scores demonstrated the intermethod as excellent ($\kappa \geq 0.864$ in all cases), and the interobserver agreement was good to excellent ($\kappa \geq 0.600$ in all cases).

Table 4 presents the diagnostic performance of the final PI-RADS category using different DWI techniques for estimating CSC. When determining CSC, the AUCs of DWI₁₀₀₀ and DWI₁₅₀₀ images were 0.710 and 0.724 for observer 1 and 0.649 and 0.656, respectively, which were not significantly different ($p = 0.11$ for observer 1; $p = 0.12$ for observer 2). The sensitivity and specificity of DWI1000 were 75.3 and 62.5% for observer 1, and 74.7 and 62.5% for observer 2, which were not significantly different compared with those of DWI1500 (78.1 and 62.5% for observer 1 and 75.3 and 62.5% for observer 2; all $p > 0.05$).

DISCUSSION

The results of the present study demonstrated excellent intermethod agreement between DWI₁₀₀₀ and DWI₁₅₀₀ at 3T in both observers regarding DWI score using the PI-RADS v2, with different experience for prostate MRI in both observers (11 years and 1 year, respectively). Furthermore, the sensitivity, specificity and AUC of DWI₁₀₀₀ were similar to those of DWI₁₅₀₀ in the detection of CSC by both observers. These findings indicate that conventional DWI₁₀₀₀ at 3T can be used for PI-RADS v2 scoring instead of higher b -value DWI₁₅₀₀. Additionally, DWI₁₀₀₀ can offer several advantages compared with DWI₁₅₀₀, such as reduced MRI examination time, improved SNR or decreased the susceptibility artefacts.

Until now, many published studies have focused on the value of PCa detection using high b -value DWI.^{4–12} However, to our knowledge, no studies have assessed the difference in DWI score using PI-RADS v2 between different higher b -value of $\geq 1000 \text{ s mm}^{-2}$ on DWI. In daily practice, radiologists perform qualitative assessment and determine the PI-RADS score of index lesions in the prostate; such evaluation may be subjective

Figure 1. A 71-year-old male with prostate cancer of Gleason score $3 + 4 = 7$ in the left PZ of the mid-gland. a, T_2 weighted image shows a 2.0-cm-sized hypointense area (arrow) in the left PZ of the mid-gland. (b–e, DW images of $b = 1000 \text{ s mm}^{-2}$ (b) and $b = 1500 \text{ s mm}^{-2}$ (c) show a focal, markedly hyperintense mass (arrow) in the corresponding site with (a), as well as a markedly hypointense signal (arrow) on ADC maps of $b = 1000$ (d) and 1500 s mm^{-2} (e). Both observers conferred DWI and final PI-RADS scores of 5 on DWI of $b = 1000$ and 1500 s mm^{-2} .

to a certain extent.¹⁵ Accordingly, it is crucial that researchers assess the variability of DWI score using the PI-RADS v2 between conventional (DWI_{1000}) and higher b -value ($b > 1000 \text{ s mm}^{-2}$) DWI. It is essential that the findings by different observers be compared in this regard.

According to the recently published PI-RADS v2 guidelines, DWI is the dominant MRI sequence for evaluating index lesions in the PZ and the secondary MRI sequence for evaluating TZ lesions. Typically, the normal PZ of the prostate gland is not fully suppressed and it continues to product a mildly increased signal

Figure 2. A 66-year-old male with prostate cancer of Gleason score $3 + 3 = 6$ in the left PZ of the mid-gland. A, T_2 weighted image shows a 0.7-cm-sized hypointense area (arrow) in the left PZ of the mid-gland. (b,C) The lesion shows a focal isointense to mild hyperintense signal (arrow) on DW images of $b = 1000 \text{ s mm}^{-2}$ (b) and a markedly hyperintense signal (arrow) on DW images of $b = 1500 \text{ s mm}^{-2}$ (c), respectively. (d, e) The corresponding ADC maps of $b = 1000$ (d) and 1500 s mm^{-2} (e) show a markedly hypointense signal (arrow). (f) On dynamic contrast-enhanced image, the lesion shows focal early enhancement (arrow). Both observers conferred a DWI score of 3 on $b = 1000$ and of 4 on $b = 1500 \text{ s mm}^{-2}$. Accordingly, the final PI-RADS score in both observers was 4 on both $b = 1000$ and 1500 s mm^{-2} .

Table 4. Diagnostic performance of final PI-RADS assessment category at different DWI for estimating CSC

PI-RADS		Observer 1			Observer 2		
		Sensitivity (%)	Specificity (%)	AUC	Sensitivity (%)	Specificity (%)	AUC
DWI ₁₀₀₀	All	75.3 (67.5–82.1)	62.5 (24.5–91.5)	0.710 (0.632–0.780)	74.7 (66.8–81.5)	62.5 (24.5–91.5)	0.649 (0.568–0.724)
	PZ	74.6 (64.6–82.4)	60.0 (14.7–94.7)	0.666 (0.569–0.756)	75.2 (65.7–83.3)	60.0 (14.7–94.7)	0.659 (0.561–0.749)
	TZ	77.8 (62.9–88.8)	66.7 (9.4–99.2)	0.796 (0.655–0.890)	66.7 (51.0–80.0)	66.7 (9.4–99.2)	0.6111 (0.460–0.748)
DWI ₁₅₀₀	All	78.1 (70.5–84.5)	62.5 (24.5–91.5)	0.724 (0.647–0.793)	75.3 (67.5–82.1)	62.5 (24.5–91.5)	0.656 (0.576–0.731)
	PZ	78.2 (68.9–85.8)	60.0 (14.7–94.7)	0.683 (0.586–0.770)	78.2 (68.9–85.8)	60.0 (14.7–94.7)	0.656 (0.558–0.746)
	TZ	77.8 (62.9–88.8)	66.7 (9.4–99.2)	0.796 (0.655–0.899)	66.7 (51.0–80.0)	66.7 (9.4–99.2)	0.6111 (0.460–0.748)
P value ^a	All	0.25	>0.999	0.11	>0.999	>0.999	0.12
	PZ	0.25	>0.999	>0.999	>0.999	>0.999	0.95
	TZ	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999

Note: AUC, area under the curve; CI, confidence interval; CSC, clinically significant cancer; DWI, diffusion-weighted imaging; PI-RADS, Prostate Imaging Reporting and Data System; PZ, peripheral zone; TZ, transition zone.

Data in parentheses are 95% confidence interval.

^ap value: comparison between DWI₁₀₀₀ and DWI₁₅₀₀.

on DWI₁₀₀₀ due to T2-shine through effects; this can obscure some tumours on DWI of $b =$ approximately 800–1000 s mm⁻². Therefore, to eliminate the T2-shine through effect in the prostate gland, the ADC or exponential map must be reviewed with high b -value DWI in evaluating PCa.¹⁶ However, this needs longer acquisition time and reduces the SNR with potential image distortions that reduce anatomic clarity and visualization of the normal landmarks of prostate.⁴

Numerous studies have been performed to determine the optimal b -value on prostate DWI.^{4–12} Several studies have compared b -value of 1000 s mm⁻² with b -value of 2000 s mm⁻² at 3T DWI; they have found that b -value of 2000 s mm⁻² was superior to b -value of 1000 s mm⁻² for PCa detection.^{5,6,11,17} Similarly, Wang et al⁷ reported b -value of 1500 s mm⁻² was more effective than b -value of 1000 or 2000 s mm⁻². Rosenkrantz et al⁴ compared different b -values (1000–5000 s mm⁻²) and reported that 1500–2500 s mm⁻² was the optimal range for PCa detection. In contrast, Koo et al¹⁷ reported that the AUC of ADC map of $b = 1000$ s mm⁻² in the experienced observer was superior to that of ADC map of $b = 2000$ s mm⁻² in detecting PCa, while the AUCs of ADC maps between $b = 1000$ and 2000 s mm⁻² were similar in the less-experienced observer. However, in their study,¹⁷ qualitative imaging assessment was performed on the basis of ADC map instead of index DWI in detecting PCa, which were in line with other recent studies^{5,8} in that the ADC derived from different b -values did not differ in terms of PCa detection, even although DWI using b -values of 1500 or 2000 s mm⁻² had higher sensitivity. These findings suggest that the limitations of DWI₁₀₀₀ in detecting PCa can be overcome using

the ADC maps of $b = 1000$ s mm⁻², which remove T2-shine through effects. However, a multicenter study with large population will be needed.

In the present study, using DWI₁₀₀₀ and DWI₁₅₀₀, the interobserver agreement of the DWI and final PI-RADS scores were excellent ($\kappa \geq 0.870$) and good to excellent ($\kappa \geq 0.771$), respectively. These results were better than those of previous reports for PI-RADS v2.^{2,3} For example, Rosenkrantz et al² reported a moderate to good interobserver agreement ($\kappa = 0.535$ –0.619) on DWI of $b = 1500$ s mm⁻² and Muller et al³ reported a similar interobserver agreement ($\kappa = 0.40$) on DWI of $b = 2000$ s mm⁻². Several potential reasons for these discrepancies are related to methodological differences. Specifically, a number of previous studies have assessed the variability of observers with various levels of experience³ or experience observers from different academic centers,² with lesion-based analysis. In contrast, our study evaluated the variability between an experienced and a less-experienced observer from a single center and they performed patient-based analysis. Another recent study of patient-based analysis performed reported that the interobserver agreement for PI-RADS score ≥ 4 was excellent ($\kappa = 0.801$). The present study corroborated these results. In addition, potential variations in DWI image quality between different MRI vendors, or different acquisition parameters, might affect the DWI score and final PI-RADS score.

In the present study, ROC curve analysis demonstrated that with a threshold score of 4, the AUCs of the experienced and less-experienced observers in estimating CSC were 0.708–0.711 and 0.649–0.656, respectively. These results are

comparable with those of several recent studies,^{6,10,11,18} even though different study designs and heterogeneous populations have been used. Regarding the PI-RADS v2 score, recent studies have reported that the AUC for detecting CSC is 0.79–0.86 for the final PI-RADS score ≥ 3 or ≥ 4 .^{3,19} However, our results showed slightly lower performance than previous studies,^{3,19} possibly because we used a different methodology; for instance, the subject of study may have differed (lesion-based vs patient-based analysis), or we may have used different criteria to define CSC.

There were several limitations to the present study. First, it was a retrospective study with only a small number of clinically insignificant cancers ($n = 8$). Patients with clinically insignificant cancer may not have undergone surgery; thus, they may have been excluded for the present study. This would introduce a potential selection bias. Second, regarding interobserver variability, only two observers from a single center assessed the index lesion. A further large study including more observers or

institutions should be performed using lesion-based or sector-based analysis. Finally, we did not use quantitative analysis to supplement the PI-RADS score on DWI. Quantitative analysis combined with PI-RADS score might improve the diagnostic performance in detecting CSC. A recent study reported that quantitative assessment, such as ADC ratio, improves the reliability of DWI scores ≥ 4 in PI-RADS v2.²⁰ However, a prospective study may be required to evaluate it.

In conclusion, DWI₁₀₀₀ at 3T demonstrates excellent agreement with DWI₁₅₀₀ in both observers for PI-RADS v2 scoring in evaluating PCa, with excellent interobserver agreement. Therefore, we believe that conventional DWI using a b -value of 1000 s mm⁻² may be sufficient to evaluate PCa.

FUNDING

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2014R1A1A2058297).

REFERENCES

- Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS prostate imaging - reporting and data system: 2015, version 2. *Eur Urol* 2016; **69**: 16–40. doi: <https://doi.org/10.1016/j.eururo.2015.08.052>
- Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, et al. Interobserver reproducibility of the PI-RADS version 2 lexicon: a Multicenter study of six experienced prostate radiologists. *Radiology* 2016; **280**: 793–804. doi: <https://doi.org/10.1148/radiol.2016152542>
- Muller BG, Shih JH, Sankineni S, Marko J, Rais-Bahrami S, George AK, et al. Prostate cancer: interobserver agreement and accuracy with the revised prostate imaging reporting and data system at multiparametric MR imaging. *Radiology* 2015; **277**: 741–50. doi: <https://doi.org/10.1148/radiol.2015142818>
- Rosenkrantz AB, Parikh N, Kierans AS, Kong MX, Babb JS, Taneja SS, et al. Prostate cancer detection using computed very high b-value diffusion-weighted imaging: how high should we go? *Acad Radiol* 2016; **23**: 704–11. doi: <https://doi.org/10.1016/j.acra.2016.02.003>
- Rosenkrantz AB, Hindman N, Lim RP, Das K, Babb JS, Mussi TC, et al. Diffusion-weighted imaging of the prostate: comparison of b1000 and b2000 image sets for index lesion detection. *J Magn Reson Imaging* 2013; **38**: 694–700. doi: <https://doi.org/10.1002/jmri.24016>
- Ueno Y, Kitajima K, Sugimura K, Kawakami F, Miyake H, Obara M, et al. Ultra-high b-value diffusion-weighted MRI for the detection of prostate cancer with 3-T MRI. *J Magn Reson Imaging* 2013; **38**: 154–60. doi: <https://doi.org/10.1002/jmri.23953>
- Wang X, Qian Y, Liu B, Cao L, Fan Y, Zhang JJ, et al. High-b-value diffusion-weighted MRI for the detection of prostate cancer at 3T. *Clin Radiol* 2014; **69**: 1165–70. doi: <https://doi.org/10.1016/j.crad.2014.07.013>
- Wetter A, Nensa F, Lipponer C, Guberina N, Olbricht T, Schenck M, et al. High and ultra-high b-value diffusion-weighted imaging in prostate cancer: a quantitative analysis. *Acta Radiol* 2015; **56**: 1009–15. doi: <https://doi.org/10.1177/0284185114547900>
- Metens T, Miranda D, Absil J, Matos C. What is the optimal b value in diffusion-weighted MR imaging to depict prostate cancer at 3T? *Eur Radiol* 2012; **22**: 703–9. doi: <https://doi.org/10.1007/s00330-011-2298-9>
- Agarwal HK, Mertan FV, Sankineni S, Bernardo M, Senegas J, Keupp J, et al. Optimal high b-value for diffusion weighted MRI in diagnosing high risk prostate cancers in the peripheral zone. *J Magn Reson Imaging* 2017; **45**. doi: <https://doi.org/10.1002/jmri.25353>
- Ohgiya Y, Suyama J, Seino N, Hashizume T, Kawahara M, Sai S, et al. Diagnostic accuracy of ultra-high-b-value 3.0-T diffusion-weighted MR imaging for detection of prostate cancer. *Clin Imaging* 2012; **36**: 526–31. doi: <https://doi.org/10.1016/j.clinimag.2011.11.016>
- Rosenkrantz AB, Chandarana H, Hindman N, Deng FM, Babb JS, Taneja SS, et al. Computed diffusion-weighted imaging of the prostate at 3 T: impact on image quality and tumour detection. *Eur Radiol* 2013; **23**: 3170–7. doi: <https://doi.org/10.1007/s00330-013-2917-8>
- Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. European Society of Urogenital Radiology ESUR prostate MR guidelines 2012. *Eur Radiol* 2012; **22**: 746–57. doi: <https://doi.org/10.1007/s00330-011-2377-y>
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74. doi: <https://doi.org/10.2307/2529310>
- Barentsz JO, Weinreb JC, Verma S, Thoeny HC, Tempny CM, Shtern F, et al. Synopsis of the PI-RADS v2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. *Eur Urol* 2016; **69**: 41–9. doi: <https://doi.org/10.1016/j.eururo.2015.08.038>
- Kim CK, Park BK, Kim B. Diffusion-weighted MRI at 3 T for the evaluation of prostate cancer. *AJR Am J Roentgenol* 2010; **194**: 1461–9. doi: <https://doi.org/10.2214/AJR.09.3654>
- Koo JH, Kim CK, Choi D, Park BK, Kwon GY, Kim B. Diffusion-weighted magnetic resonance imaging for the evaluation of prostate cancer: optimal B value at 3T. *Korean J Radiol* 2013; **14**: 61–9. doi: <https://doi.org/10.3348/kjr.2013.14.1.61>

18. de Perrot T, Scheffler M, Boto J, Delattre BM, Combescure C, Pustaszeri M, et al. Diffusion in prostate cancer detection on a 3T scanner: How many b-values are needed? *J Magn Reson Imaging* 2016; **44**: 601–9. doi: <https://doi.org/10.1002/jmri.25206>
19. Park SY, Jung DC, Oh YT, Cho NH, Choi YD, Rha KH, et al. Prostate Cancer: pi-rads version 2 helps preoperatively predict clinically Significant cancers. *Radiology* 2016; **280**: 108–16. doi: <https://doi.org/10.1148/radiol.16151133>
20. Park SY, Shin SJ, Jung DC, Cho NH, Choi YD, Rha KH, et al. PI-RADS version 2: quantitative analysis aids reliable interpretation of diffusion-weighted imaging for prostate cancer. *Eur Radiol* 2017; **27**. doi: <https://doi.org/10.1007/s00330-016-4678-7>