Comparison of quantitative apparent diffusion coefficient parameters with prostate imaging reporting and data system V2 assessment for detection of clinically significant peripheral zone prostate cancer

Elmira Hassanzadeh,[1,2] Francesco Alessandrino[1,3] Olutayo I. Olubiyi,[1,4] Daniel I. Glazer,[1] Robert V. Mulkern,[5] Andriy Fedorov,[1] Clare M. Tempany,[1] Fiona M. Fennessy[1,3]

[1] Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA
[2] Department of Surgery, University of Illinois at Chicago, 1200 W Harrison St, Chicago, IL 60607, USA
[3] Department of Imaging, Dana Farber Cancer Institute, Boston, MA, USA
[4] Department of Radiology, Mercy Catholic Medical Center, 1500 Lansdowne Avenue, Darby, PA, USA
[5] Department of Radiology, Boston Children's Hospital, Harvard Medical School, 300 Longwood Ave, Boston, MA, USA

Abstract

Purpose: To compare diagnostic performance of PI-RADSv2 with ADC parameters to identify clinically significant prostate cancer (csPC) and to determine the impact of csPC definitions on diagnostic performance of ADC and PI-RADSv2.

Methods: We retrospectively identified treatment-naïve pathology-proven peripheral zone PC patients who underwent 3T prostate MRI, using high b-value diffusion-weighted imaging from 2011 to 2015. Using 3D slicer, areas of suspected tumor (T) and normal tissue (N) on ADC ($b = 0, 1400$) were outlined volumetrically. Mean ADC_T, mean ADC_N, ADC_{ratio} (ADC_T/ADC_N) were calculated. PI-RADSv2 was assigned. Three csPC definitions were used: (A) Gleason score (GS) $\geq 4 + 3$; (B) GS $\geq 3 + 4$; (C) MRI-based tumor volume > 0.5 cc. Performances of ADC parameters and PI-RADSv2 in identifying csPC were measured using nonparametric comparison of receiver operating characteristic curves using the area under the curve (AUC).

Results: Eighty five cases met eligibility requirements. Diagnostic performances (AUC) in identifying csPC using three definitions were: (A) ADC_T (0.83) was higher than PI-RADSv2 (0.65, $p = 0.006$); (B) ADC_T (0.86) was higher than ADC_{ratio} (0.68, $p < 0.001$), and PI-RADSv2 (0.70, $p = 0.04$); (C) PI-RADSv2 (0.73) performed better than ADC_{ratio} (0.56, $p = 0.02$). ADC_T performance was higher when csPC was defined by A or B versus C ($p = 0.038$ and $p = 0.01$, respectively). ADC_{ratio} performed better when csPC was defined by A versus C ($p = 0.01$). PI-RADSv2 performance was not affected by csPC definition.

Conclusions: When csPC was defined by GS, ADC parameters provided better csPC discrimination than PI-RADSv2, with ADC_T providing best result. When csPC was defined by MRI-calculated volume, PI-RADSv2 provided better discrimination than ADC_{ratio}. csPC definition did not affect PI-RADSv2 diagnostic performance.

Prostate cancer (PC) is the most common malignancy other than skin cancer and the second most lethal cancer in American men, with more than 3.3 million men currently living with PC in the United States [1]. The 5-year

Correspondence to: Francesco Alessandrino; *email:* falessandrino@bwh.harvard.edu

survival rate ranges from 29.3% in patients with metastatic PC to 100% in patients with localized disease [1, 2]. In addition, trends in PC diagnosis and mortality rates in the last 30 years have shown that a significant increase in early diagnosis is associated with a significant decrease in mortality rates [2, 3]. Nonetheless, the rate of over diagnosis of non-clinically significant prostate cancer (nsPC) has been shown to be as much as 67%, leading to unnecessary treatment and associated morbidity [2]. The key to optimal PC management is the detection of clinically significant cancers that require prompt treatment [3]. There are several definitions of "clinically significant prostate cancer" (csPC), such as biopsy Gleason score (GS) $\geq 3 + 4$; GS $\geq 4 + 3$; and/or pathological volume >0.5 cc [4–6]. The definition of csPC is important when we seek to determine the role and accuracy of multi-parametric magnetic resonance imaging (mpMRI) in its detection.

Apparent diffusion coefficient (ADC), derived from Diffusion-Weighted Imaging (DWI), plays an important role in mpMRI, as ADC values in men suspected of having PC have shown significant inverse correlation with GS when compared with biopsy and final pathological results [7–9]. ADC is lower in neoplastic prostatic tissue than non-neoplastic tissue, and it can differentiate low from high-grade PC [9, 10]. Importantly, the use of high b-values between 1000 and 1500 s/mm^2 in DWI is more sensitive to tumor detection, as conspicuity of PC is improved, especially in those adjacent to or invading the anterior fibromuscular stroma [11–13]. Prostate mpMRI has emerged as an excellent method to detect and localize prostate cancer and thus guide treatment selection [14, 15]. There have been many studies demonstrating the major role of DWI in detection of Gleason pattern 4 disease, which have led to an interest in investigating the role of MRI in triaging patients before prostate biopsy. Two recent studies, the multi-center prospective PROMIS study and the bi-parametric IMPROD trial have determined that MRI can reduce the biopsy rate by 27% and 24%, respectively [15, 16].

As both qualitative and quantitative imaging metrics are influenced by the underlying mpMRI acquisition parameters, a major effort to standardize mpMRI acquisition, interpretation and reporting led to formation of Prostate Imaging – Reporting and Data System (PI-RADS) in 2012 by the European society of Urogenital Radiology [17, 18]. Subsequently, a broader International group developed PI-RADS version 2 (v2), which utilizes a 5 point scale to determine the likelihood of csPC. PI-RADS v2 defines csPC on pathology/histology as having GS ≥ 7 (including $3 + 4$ with prominent but not predominant Gleason 4 component), and/or pathology volume ≥ 0.5 cc, and/or extra prostatic extension [16]. PI-RADS v2 is based upon qualitative interpretation of DWI/ADC, dynamic contrast-enhanced MRI (DCE-MRI) and T2-weighted imaging (T2WI). However, for peripheral zone (PZ) lesions, where 70% to 80% of PC is known to occur, DWI and ADC signal intensity are the dominant sequence used by PI-RADS v2 to determine likelihood of csPC [16–19]. In addition, the recommendation to use qualitative high b-value (>1400 s/mm^2) sequences has been suggested by the PI-RADS v2 committee [13].

Importantly, it recently has been shown that pathological tumor volume affects PI-RADS v2 performance, which was shown to identify only 24% of tumors GS $\geq 4 + 3$ with a pathological tumor volume of ≤ 0.5 cc [20]. This underscores the importance of including tumor volume in the assessment of diagnostic performance of qualitative and quantitative imaging metrics.

The purpose of our study was twofold: first, to compare the performance of qualitative assessment with PI-RADS v2 versus quantitative volumetric ADC assessment in detection of csPC in the PZ; and second, to determine the effect of multiple definitions of csPC on the diagnostic performance of ADC and PI-RADS v2.

Materials and methods

Patients

This retrospective study was HIPAA compliant and approved by the local Institutional Review Board. With the help of our institutional search engine, we identified patients who underwent endorectal coil mpMRI at 3T between January 2011- January 2015, were treatment naïve, and had pathologically proven PC within 6 months of mpMRI. Prior to the radiologist reviewing the cases, the mpMRIs underwent pre-screening to exclude cases that did not have high b-value (at least 0, 1400 s/mm^2) DWI imaging.

MRI technique

All MRI examinations were performed on a GE Signa HDx 3.0 T magnet (GE Healthcare, Waukesha, WI) using an eight-channel abdominal array and endorectal coil (Medrad, Pittsburgh, PA), as previously described [21]. T1-weighted images were generated from spoiled gradient echo (SPGR) sequences with repetition time (TR)/echo time (TE)/α = 385/6.2 ms/65°, over a 16-cm^2 field of view (FOV), 384 × 192 resolution and 3 mm slice thickness, T2-weighted images from fast relaxation fast spin echo (FRFSE) sequence, TR/TE = 3500 ms/ 102 ms over a 16-cm^2 FOV, with 384 × 224 resolution and 3 mm slice thickness, and DWI images from single-shot echo planar imaging (EPI) sequences, TR/ TE = 2500 ms/65 ms with b-values of 0, 500 and 1400 s/mm^2 were acquired with 128 × 96 resolution and 3 mm slice thickness. DCE-MRI was performed with a 3D SPGR sequence with full-gland coverage. Gadopentetate dimeglumine (Magnevist, Berlex Laboratories, Wayne,

NJ) was injected intravenously (0.15 mmol/kg; rate 3 mL/s).

Clinical and histopathologic information collection

Age and serum prostate specific antigen (PSA) at the time of PC diagnosis were collected using our electronic medical records database. We retrospectively reviewed the histopathology reports of the identified patients and recorded the primary and secondary GS. The pathology specimens were either from prostatectomy, MRI guided biopsy (MRGBx), or transrectal ultrasound guided prostate biopsy (TRUS-Bx). All prostate specimens were processed according to the Stanford protocol [22].
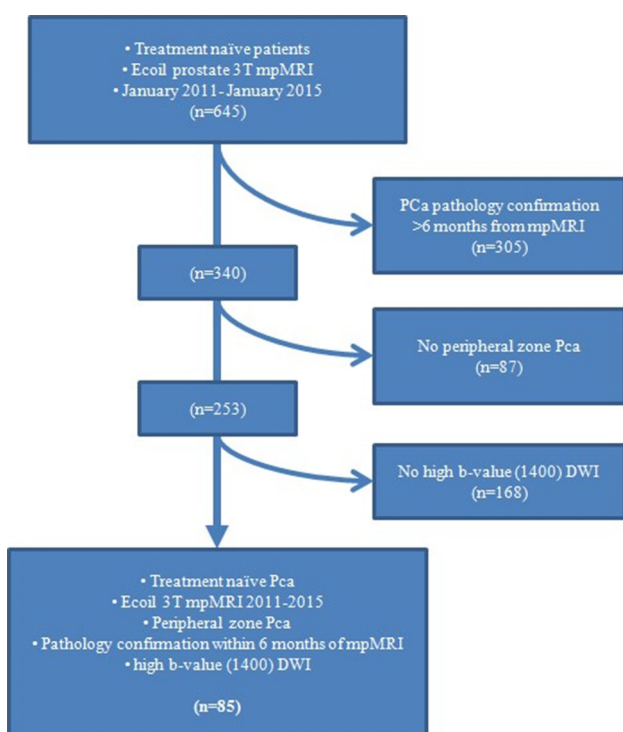

Fig. 1. Flowchart of patient enrollment, exclusion criteria, and final study population. PCa, prostate cancer; mpMRI, multiparametric magnetic resonance imaging; DWI, diffusion-weighted imaging; ecoil, endorectal coil.

Image analysis

A single radiologist (F.M.F.), with over 10 years of experience interpreting prostate MRI, blinded to GS, reviewed the de-identified mpMRIs of each patient to identify cases with PZ index lesions. Tumor (T) and normal (N) PZ tissue volumes of interest were then outlined on ADC (b 0, 1400 s/mm^2) maps using 3D Slicer, an open source image processing software (www.slicer.org) which has shown high reproducibility for mpMRI-based quantitative measurements [23, 24]. MRI index lesion volume was determined from ADC maps, and ADC_{mean} for T (ADC_T) and N (ADC_N) were all calculated on 3D Slicer. ADC_{ratio} was defined as ADC_T/ADC_N. In the same sitting, an overall PI-RADS v2 assessment category (based upon T2, DCE, and DWI/ADC) was assigned to T [13]. PI-RADS v2 overall assessment category ≥ 4 were defined as "high" risk for csPC.

Definition of clinically significant PC

PZ Tumors were categorized as csPC based on either histopathology or tumor volume, using the following three definitions: Method A: index tumors with GS $\geq 4 + 3$; Method B: index tumors with GS $\geq 3 + 4$; Method C: MRI-based tumor volume ≥ 0.5 cc, calculated volumetrically on multiple slices of ADC maps, based upon the premise that pathologic-based tumor volume ≥ 0.5 cc is csPC [4–6].

Statistical analysis

Continuous variables were summarized using means and ranges, and categorical variables were summarized using frequencies and percentages. Each ADC metric (ADC_T, ADC_N, and ADC_{ratio}) was compared between csPC and nsPC groups for each method (A, B, C) using a two-sample T test. Tumor ADC metrics (ADC_T and ADC_{ratio}) were correlated with GS using the Pearson test. Fisher's Exact test was used to evaluate for associations between high PI-RADS v2 (≥ 4) and csPC, as defined by Methods A, B, and C.

Receiver operator characteristic (ROC) curves were plotted for PI-RADS v2, ADC_T, and ADC_{ratio} to evaluate their performance in detecting csPC, and perfor-

Table 1. Break down of Gleason scores based on reference pathology

Gleason score	Pathology reference standard				
	Prostatectomy	TRUS-Bx	MRGBx	Total	%
3 + 3	6	11	1	18	21.2
3 + 4	30	3	0	33	38.8
4 + 3	12	2	0	14	16.5
4 + 4	3	3	3	9	10.6
4 + 5	3	7	1	11	12.9
	54	26	5	85	100

TRUS-Bx, transrectal ultrasound guided biopsy; MRGBx, MRI guided biopsy

Table 2. Distribution of lesions according to clinically significant prostate cancer definitions

PC definition	csPC	nsPC	Total lesions
Method A (GS ≥ 4 + 3)	34	51	85
Method B (GS ≥ 3 + 4)	67	18	
Method C (Tvol > 0.5 cc)	54	31	

PC, prostate cancer; csPC, clinically significant prostate cancer; nsPC, non-clinically significant prostate cancer; GS, Gleason score; Tvol, tumor volume

Table 3. Comparison of ADC parameters between significant and non significant peripheral zone prostate cancer using three different definitions of clinical significance (A: GS ≥ 4 + 3, B: GS ≥ 3 + 4, C: volume >0.5 cc)

	ADC_T			ADC_N			ADC_{ratio}		
	Mean ± SD		p value	Mean ± SD		p value	Mean ± SD		p value
	csPC	nsPC		csPC	nsPC		csPC	nsPC	
A	809.5 ± 143.1	1027.6 ± 204.8	$p < 0.001$	1365.3 ± 225.4	1479 ± 255.7	$p = 0.03$	0.60 ± 0.09	0.71 ± 0.14	$p < 0.001$
B	881.4 ± 185.6	1147.6 ± 162.8	$p < 0.001$	1380.4 ± 227.2	1624.9 ± 235.7	$p < 0.001$	0.65 ± 0.07	0.71 ± 0.14	$p = 0.01$
C	901.5 ± 189.6	1001.1 ± 233.0	$p = 0.035$	1376.7 ± 215.5	1528.9 ± 275.8	$p = 0.006$	0.66 ± 0.12	0.67 ± 0.15	$p = 0.79$

csPC, clinically significant prostate cancer; nsPC, non-clinically significant prostate cancer

Table 4. Performance of ADC and PI-RADS v2 parameters in differentiating clinically significant from non-clinically significant prostate cancer

	AUC	Cutoff	Correctly classified %	Sensitivity	Specificity
A. Clinical significance = Gleason score ≥ 4 + 3					
1. ADC_T	0.83	1010	70.59	94.29	54
2. ADC_{ratio}	0.77	0.68	70.59	88.57	58
3. PI-RADSv2	0.65	(≥ 4)	48.24	100	12
B. Clinical significance = Gleason score ≥ 3 + 4					
4. ADC_T	0.86	1069.13	83.53	88.06	66.67
5. ADC_{ratio}	0.70	0.68	69.41	70.15	66.67
6. PI-RADSv2	0.70	(≥ 4)	81.18	97.01	22.22
C. Clinical significance = MRI-based tumor volume >0.5 cc					
7. ADC_T	0.66	1069.13	68.24	85.19	38.71
8. ADC_{ratio}	0.56	0.68	64.71	70.37	54.84
9. PI-RADSv2	0.72	(≥ 4)	65.88	96.30	12.90

Significant difference observed between: parameter comparisons in method A: ADC_T vs PI-RADSv2, $p < 0.01$, parameter comparisons in method B: ADC_T vs ADC_{ratio} $p < 0.01$; ADC_T vs PI-RADSv2, $p = 0.04$, parameter comparisons in method C: ADC_T vs ADC_{ratio}, $p = 0.04$; ADC_{ratio} vs PI-RADSv2, $p = 0.02$

mance statistics were calculated (AUC, sensitivity, specificity, PPV, NPV). The AUC of all ROC plots were compared using a nonparametric approach, previously described by DeLong et al. [25]. This nonparametric approach was also used to compare the diagnostic performances (ROC for differentiating csPC from nsPC) of each ADC metric and overall PI-RADS v2 assessment category among the three definitions of csPC.

A two-tailed p value less than 0.05 was considered significant. All statistical analyses were performed using STATA (Version 11.2 StataCorp College Station, Texas USA).

Results

Patients and tumor characteristics

A total of 645 patients were identified. Using our institutional medical record search engine, we excluded those who did not have prostate pathology confirmation within a 6-month window of the mpMRI ($n = 305$), did not have an index PZ tumor on pathology ($n = 87$), or did not undergo high b-value (1400 s/mm^2) DWI sequences ($n = 168$) to arrive at a final population of 85 patients (Fig. 1). Their mean was age 61.2 years [range 43–76 years] and mean PSA 6.52 ng/ml [range 0.01–55.88 ng/ml].

The pathologic diagnosis of PC was based on prostatectomy ($n = 54$), TRUS-Bx ($n = 26$), and MRGBx ($n = 5$) (Table 1). Those who underwent TRUS-Bx all had a repeat TRUS-Bx, the results of which were in agreement with the first TRUS-Bx pathology. The median GS was 7 [range 6–9]. The distribution of GS was as follows: GS = 6 ($n = 18$); GS = 7 (47); GS = 8 ($n = 9$); and GS = 9 ($n = 11$). In patients with GS 7, primary Gleason grade was 3 (3 + 4) in 33 patients and 4 (4 + 3) in 14 patients. Based on Method A, 34/85 patients had csPC; based on Method B, 67/85 had csPC, and based on Method C,

54/85 patients had csPC. Distribution of lesions according to csPC definitions are reported in Table 2.

ADC in differentiating csPC from nsPC using three different methodologies for defining csPC

ADC_T was significantly lower in those with csPC compared to nsPC, regardless of the definition of csPC ($p < 0.001$, $p < 0.001$, $p = 0.035$ for Methods A, B, and C, respectively) (Table 3). ADC_N was also significantly lower in those with csPC compared to nsPC ($p < 0.03$, $p < 0.001$; $p = 0.006$ for Methods A, B, and C, respectively) (Table 3). ADC_{ratio} was significantly lower in those with csPC compared to nsPC when csPC was defined with Method A and B ($p < 0.001$ and $p = 0.01$, respectively), but not with Method C. There was a negative correlation between ADC_T and GS (ρ: −0.51, $p < 0.001$), and between ADC_{ratio} and GS (ρ: −0.39, $p < 0.001$).

PI-RADS v2 in differentiating csPC from nsPC using three different methodologies for defining csPC

There was a significant association between high PI-RADS v2 assessment category and a diagnosis of csPC using Method A ($p = 0.04$) and Method B ($p < 0.01$). There was no association between high PI-RADS v2 and a csPC diagnosis when csPC was defined according to Method C ($p = 0.11$).

Comparison of diagnostic performance of ADC and PI-RADS v2 in detecting csPC

Using the three methods of defining csPC, the ROC curves for ADC metrics and PI-RADS v2 assessment category used to differentiate csPC from nsPC were plotted. Performance characteristics (AUC, sensitivity, specificity, NPV, PPV) of each metric (ADC_T, ADC_{ratio}, and PI-RADS v2) are presented in Table 4.

Method A (GS ≥ 4 + 3): The AUC of ADC_T was 0.83 for a cutoff of 1010 mm²/s; AUC of ADC_{ratio} was 0.77 for a cutoff of 0.68; the AUC of PI-RADS v2 was 0.65. The AUC of ADC_T (0.83) was higher than PI-RADS v2 (0.65) ($p = 0.006$). There was no difference in AUC between ADC_T and ADC_{ratio}, or between ADC_{ratio} and PI-RADS v2.

Method B (GS ≥ 3 + 4): The AUC of ADC_T was 0.86 for a cutoff of 1069.13 mm²/s; the AUC of ADC_{ratio} was 0.70 for a cutoff of 0.68; the AUC of PI-RADS v2 was 0.65. The AUC of ADC_T (0.86) was higher than that of ADC_{ratio} (0.70) ($p < 0.001$) and PI-RADS v2 (0.70) ($p = 0.04$). There was no difference in AUC between ADC_{ratio} and PI-RADS v2.

Method C (MRI-based tumor volume > 0.5 cc): The AUC of ADC_T was 0.66 for a cutoff of 1069.13 mm²/s; the AUC of ADC_{ratio} was 0.70 for a cutoff of 0.68; the AUC of PI-RADS v2 was 0.65. The AUC of PI-RADS v2 (0.72) was higher than AUC of ADC_{ratio} (0.56) ($p = 0.02$). AUC of ADC_T (0.66) was also higher than AUC of ADC_{ratio} ($p = 0.04$).

Effect of csPC definition on the diagnostic performance of ADC and PI-RADS v2

Performance of ADC_T was better when csPC was defined pathologically (Methods A and B), compared with volumetrically (Method C) (A vs C, $p = 0.038$; B vs C, $p = 0.011$). ADC_{ratio} performed better when csPC was defined according to Method A, compared with Method C ($p = 0.01$), with no difference observed between performance of ADC_{ratio} for method A compared to B, or B compared to C. Performance of PI-RADS v2 did not differ when csPC was defined by methods A, B, or C.

Discussion

This study was motivated by the introduction of PI-RADS v2, whose aim is to standardize acquisition, interpretation, and reporting of mpMRI of the prostate in order to improve the accuracy of mpMRI in the diagnosis of csPC. In this study, we demonstrated that index tumor ADC_T outperforms PI-RADS v2 in detection of csPC in the PZ when we utilize a pathological GS definition of clinical significance, irrespective of whether this is defined as GS ≥ 3 + 4 or GS ≥ 4 + 3. Considering that ADC metrics have been shown to correlate with tumor GS, this finding is not unexpected.

However, if we consider csPC to be an index tumor with an MRI-based volume of ≥ 0.5 cc (Method C), ADC_T did not perform better than PI-RADS v2 in detecting csPC. As a recent retrospective analysis showed, tumors measuring < 5 mm diameter on pathology are much less likely to be of higher Gleason grade [26]. It is therefore understandable that smaller contoured tumors may have a higher ADC score due to volume averaging with adjacent normal tissue, leading to a lower PI-RADS v2 score. Conversely, larger lesions are more likely to be given a higher PI-RADS score [16, 27]. Vargas et al. have also evaluated PI-RADS v2 accuracy in detecting csPC defined as tumor volume > 0.5 cc or GS ≥ 4 + 3 [20]. While they found PI-RADS v2 helpful in detecting lesions > 0.5 cc, they also noted its limitation in assessing high GS lesions which were < 0.5 cc.

It is important to note that our definition of csPC being ≥ 0.5 cc tumor volume in this study was based on volumetric MR image contouring of tumor on ADC, and not on pathologically-derived tumor volume from the prostatectomy specimen. However, MRI is known to significantly underestimate lesion size compared to whole mount prostatectomy analysis, particularly so for ADC [28, 29]. It is therefore possible that volumetric contouring of tumor volume based upon ADC underesti-

mates the actual number of tumors that would be considered csPC if it was based on pathological tumor volume definition. As such, if we underestimate the number of csPC cases based on MRI-calculated volume, it may be expected that performance of ADC metrics and PI-RADS v2 are not as good for the volumetric definition of csPC (Method C), compared to those based upon GS (method A and B), which is what was we found.

In our study, PI-RADS v2 correctly classified 81.2% of GS $\geq 3 + 4$ but only 48.24% of GS $\geq 4 + 3$ (Table 4). The high sensitivity and very low specificity of PI-RADS-v2 across the three methods in Table 3 suggests that this tool is good to identify csPC in a selected population of patients with known PCa, as a negative test will be useful to rule out csPC, but it should not be used as a confirmatory test. This contrasts with a recent study where PI-RADS-v2 was found to have a sensitivity and specificity of over 80% for detection csPC [30]. The difference in performance characteristics may be explained by their use of a PI-RADS-v2 score of 3 ("the presence of clinically significant cancer is equivocal") as a cutoff point for detection of csPC, whereas we used a score of 4 ("clinically significant cancer is likely to be present") as a cutoff.

Our results comparing ADC metrics between tumor and normal tissue are similar to those previously reported, with significantly lower mean ADC_T than those in healthy-appearing PZ prostate tissue (ADC_N) [11, 31, 32]. Various ADC metrics have been investigated, including ADC_{mean} and ADC_{ratio}, for detection of csPC. ADC_{mean} values can either be calculated on a single slice that contains the largest volume of tumor, or on multiple slices to include whole tumor 3D volume [11, 31, 33]. ADC_{ratio} is defined as the ratio of ADC values of tumor to normal tissue, which aims to minimize the impact of study-specific imaging protocols [11, 34]. We also found a significant negative correlation between ADC parameters and GS, which is in agreement with previous studies that have evaluated ADC in assessment of aggressive PC [12, 32, 35]. Interestingly, we also found mean ADC_N in the PZ to be significantly lower in the normal tissue of those with csPC group compared to those with nsPC, irrespective of the definition of csPC. To our knowledge, this finding that the background of "normal" ADC values in the prostate (i.e., in areas where no visible tumor by mpMRI) are significantly different in those patients with pathologically confirmed aggressive PC elsewhere in the prostate compared to those with less aggressive tumors, has not been previously reported. We speculate that these findings may be due to field cancerization, a known phenomenon in the histopathology literature, and currently being evaluated in the prostate [36].

In a recent study, Boesen et al. reported a better performance of ADC_{ratio} compared to ADC_T in detecting high-grade cancer (defined as GS $\geq 7(4 + 3)$) [35].

Our results suggest that ADC_{ratio} and ADC_T perform equally well using a similar classification method (Method A). Another study by Woo et al. demonstrated that ADC_T and ADC_{ratio} performed equally well in predicting high-grade PC, defined as (GS $> 6 (3 + 3)$) [8]. Using a similar csPC classification in our study (Method B), we found ADC_T performed better than ADC_{ratio}. It is possible that these differences in performance of ADC metrics amongst studies may be a reflection of different methods of ADC parameter calculations (single slice, versus volumetric), and also the method of calculation and observed ADC_N scores in csPC, which is taken into consideration for calculation of ADC_{ratio}. Also, it should also be noted that we used different imaging techniques and analyses than employed by Boesen and Woo [8, 35]. For ADC map calculation, we used b-values of 0, 500, and 1400 s/mm^2 while Boesen et al. used 100 and 800 s/mm^2 [35]. The non-monoexponential nature of signal decays with b-value in prostate tissues complicates comparisons between studies using different b-value ranges [37]. However, our use of higher b-values may have had an impact on tumor conspicuity as supported by a recent study that demonstrated AUC for differentiating high risk PC from low risk PC increases with an increase in b-value, reaching maximum AUC at $b = 1600$ s/mm [38]. Another possible explanation for the differences is that we performed a volumetric analysis of ADC values, as the lesion contouring was performed on multiple ADC slices to cover the entire tumor, enabling us to evaluate heterogeneous tumors with improved accuracy. In the aforementioned studies, the authors outlined the tumor only on the single ADC slice where the largest area of the tumor was visible.

Our study has several limitations: first, due to the specific enrollment criteria of requiring a high b-value DWI scan and pathologically documented PZ PC, several cases were excluded. Second, in our final population of 85 patients, we used prostatectomy as a pathology reference for the majority of our cases (54/85), but not in all. In 26/85 (30.5%), we relied on TRUS-Bx pathology result, known to have a limitation of under-detection and under-grading csPC [39]. However, in all cases, a repeat TRUS-Bx did not result in up- or down-grading of GS. Many of these patients were on active surveillance and had a GS of $3 + 3$, thereby enabling us to evaluate the performance of the diagnostic tools in a population of patients with known PC but that did not require prostatectomy. Third, we did not calculate pathologically-derived tumor volumes for our patient cohort, and as such could not use a pathological definition of csPC ≥ 0.5 cc to compare with GS definitions. Lastly, mpMRI were analyzed by a single reader with more than 10 years of experience in reading prostate mpMRI, and interreader agreement was not evaluated. However, a recent study evaluated interreader agreement of PI-RADS v2 ≥ 4 for detection of clinically significant PZ

prostate cancer (GS ≥ 3 + 4), and showed almost perfect agreement ($k = 0.91$) [40]. Given the known substantial interreader agreement for PI-RADS v2 ≥ 4 for detection of PZ prostate cancer, we deemed the use of multiple radiologists not necessary.

In conclusion, quantitative ADC metrics outperform qualitative PI-RADS v2 in detection of csPC. When csPC was defined as MRI-based tumor volume > 0.5 cc, qualitative PI-RADS v2, and quantitative mean ADC_T provide similar discrimination. How csPC was defined did have an effect on the diagnostic performance of the ADC metrics, which performed better when csPC was defined by GS. However, the definition of csPC did not affect PI-RADS v2 performance. These findings have implications as we continue to refine PI-RADS and should be considered when planning for the next version of the guidelines (v3) towards improved assessment in the detection of csPC.

References

1. Miller KD, Siegel RL, Lin CC, et al. (2016) Cancer treatment and survivorship statistics. CA Cancer J Clin 66(4):271–289
2. Loeb S, Bjurlin MA, Nicholson J, et al. (2014) Overdiagnosis and overtreatment of prostate cancer. Eur Urol 65(6):1046–1055
3. Etzioni R, Penson DF, Legler JM, et al. (2002) Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst 94(13):981–990
4. Ahmed HU, Hu Y, Carter T, et al. (2011) Characterizing clinically significant prostate cancer using template prostate mapping biopsy. J Urol 186(2):458–464
5. Rasiah KK, Stricker PD, Haynes A-M, et al. (2003) Prognostic significance of Gleason pattern in patients with Gleason score 7 prostate carcinoma. Cancer 98(12):2560–2565
6. Wolters T, Roobol MJ, van Leeuwen PJ, et al. (2011) A critical analysis of the tumor volume threshold for clinically insignificant prostate cancer using a data set of a randomized screening trial. J Urol 185(1):121–125
7. Kumar V, Jagannathan NR, Kumar R, et al. (2007) Apparent diffusion coefficient of the prostate in men prior to biopsy: determination of a cut-off value to predict malignancy of the peripheral zone. NMR Biomed 20(5):505–511
8. Woo S, Kim SY, Cho JY, et al. (2016) Preoperative evaluation of prostate cancer aggressiveness: using ADC and ADC ratio in determining Gleason score. AJR 207(1):114–120
9. Hambrock T, Somford DM, Huisman HJ, et al. (2011) Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. Radiology 259(2):453–461
10. Vargas HA, Akin O, Franiel T, et al. (2011) Diffusion-weighted endorectal MR imaging at 3 T for prostate cancer: tumor detection and assessment of aggressiveness. Radiology 259(3):775–784
11. Metens T, Miranda D, Absil J, et al. (2012) What is the optimal b value in diffusion-weighted MR imaging to depict prostate cancer at 3T? Eur Radiol 22(3):703–709
12. Kim CK, Park BK (2010) Kim B (2010) High-b-value diffusion-weighted imaging at 3 T to detect prostate cancer: comparisons between b values of 1,000 and 2,000 s/mm2. AJR 194(1):W33–W37
13. Weinreb JC, Barentsz JO, Choyke PL, et al. (2016) PI-RADS prostate imaging—reporting and data system: 2015, version 2. Eur Urol 69(1):16–40
14. Bratan F, Niaf E, Melodelima C, et al. (2013) Influence of imaging and histological factors on prostate cancer detection and localisation on multiparametric MRI: a prospective study. Eur Radiol 23(7):2019–2029
15. Jambor I, Boström PJ, Taimen P, et al. (2017) Novel biparametric MRI and targeted biopsy improves risk stratification in men with a clinical suspicion of prostate cancer (IMPROD Trial). J Magn Reson Imaging. doi:10.1002/jmri.25641
16. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. (2017) Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. Lancet 389(10071):815–822
17. Kozlowski P, Chang SD, Goldenberg SL (2008) Diffusion-weighted MRI in prostate cancer—comparison between single-shot fast spin echo and echo planar imaging sequences. Magn Reson Imaging 26(1):72–76
18. Barentsz JO, Richenberg J, Clements R, et al. (2012) ESUR prostate MR guidelines 2012. Eur Radiol 22(4):746–757
19. McNeal JE, Redwine EA, Freiha FS, et al. (1988) Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. Am J Surg Pathol 12(12):897–906
20. Vargas HA, Hötker AM, Goldman DA, et al. (2016) Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference. Eur Radiol 26(6):1606–1612
21. Hegde JV, Mulkern RV, Panych LP, et al. (2013) Multiparametric MRI of prostate cancer: an update on state-of-the-art techniques and their performance in detecting and localizing prostate cancer. J Magn Reson Imaging 37(5):1035–1054
22. Srigley JR, Humphrey PA, Amin MB, et al. (2009) Protocol for the examination of specimens from patients with carcinoma of the prostate gland. Arch Pathol Lab Med 133(10):1568–1576
23. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. (2012) 3D Slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging 30(9):1323–1341
24. Fedorov A, Vangel MG, Tempany CM, Fennessy FM (2017) Multiparametric magnetic resonance imaging of the prostate: repeatability of volume and apparent diffusion coefficient quantification. Investig Radiol 52(9):538–546
25. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44(3):837–845
26. Stensland KD, Coutinho K, Hobbs AR, et al. (2016) Are magnetic resonance imaging undetectable prostate tumours clinically significant? Results of histopathological analyses. Arab J Urol 14(4):256–261
27. Truong M, Hollenberg G, Weinberg E, et al. (2017) Impact of Gleason subtype on prostate cancer detection using multiparametric MRI: correlation with final histopathology. J Urol. doi:10.1016/j.juro.2017.01.077
28. Turkbey B, Mani H, Aras O, et al. (2012) Correlation of magnetic resonance imaging tumor volume with histopathology. J Urol 188(4):1157–1163
29. Priester A, Natarajan S, Khoshnoodi P, et al. (2017) Magnetic resonance imaging underestimation of prostate cancer geometry: use of patient specific molds to correlate images with whole mount pathology. J Urol 197(2):320–326
30. Zhao C, Gao G, Fang D, et al. (2016) The efficiency of multi-parametric magnetic resonance imaging (mpMRI) using PI-RADS version 2 in the diagnosis of clinically significant prostate cancer. Clin Imaging 40(5):885–888

31. Lin WC, Westphalen AC, Silva GE, et al. (2016) Comparison of PI-RADS 2, ADC histogram-derived parameters, and their combination for the diagnosis of peripheral zone prostate cancer. Abdom Radiol (NY) 41(11):2209–2217

32. Glazer DI, Hassanzadeh E, Fedorov A, et al. (2017) Diffusion-weighted endorectal MR imaging at 3T for prostate cancer: correlation with tumor cell density and percentage Gleason pattern on whole mount pathology. Abdom Radiol (NY) 42(3):918–925

33. Donati OF, Mazaheri Y, Afaq A, et al. (2014) Prostate cancer aggressiveness: assessment with whole-lesion histogram analysis of the apparent diffusion coefficient. Radiology 271(1):143–152

34. Lebovici A, Sfrangeu SA, Feier D, et al. (2014) Evaluation of the normal-to-diseased apparent diffusion coefficient ratio as an indicator of prostate cancer aggressiveness. BMC Med Imaging 14:15

35. Boesen L, Chabanova E, Løgager V, et al. (2015) Apparent diffusion coefficient ratio correlates significantly with prostate cancer gleason score at final pathology. J Magn Reson Imaging 42(2):446–453

36. Jones AC, Antillon KS, Jenkins SM, et al. (2015) Prostate field cancerization: deregulated expression of macrophage inhibitory cytokine 1 (MIC-1) and platelet derived growth factor A (PDGF-A) in tumor adjacent tissue. PLoS One 10(3):e0119314

37. Mulkern RV, Barnes AS, Haker SJ, et al. (2006) Biexponential characterization of prostate tissue water diffusion decay curves over an extended b-factor range. Magn Reson Imaging 24(5):563–568

38. Agarwal HK, Mertan FV, Sankineni S, et al. (2017) Optimal high b-value for diffusion weighted MRI in diagnosing high risk prostate cancers in the peripheral zone. J Magn Reson Imaging 45(1):125–131

39. Ukimura O, Coleman JA, de la Taille A, et al. (2013) Contemporary role of systematic prostate biopsies: indications, techniques, and implications for patient care. Eur Urol 63(2):214–230

40. Purysko AS, Bittencourt LK, Bullen JA, et al. (2017) Accuracy and interobserver agreement for prostate imaging reporting and data system, version 2, for the characterization of lesions identified on multiparametric MRI of the prostate. AJR 209(2):339–349