



# Biparametric prostate MRI: impact of a deep learning-based software and of quantitative ADC values on the inter-reader agreement of experienced and inexperienced readers

Stefano Cipollari<sup>1</sup> · Martina Pecoraro<sup>1</sup> · Ali Forookhi<sup>1</sup> · Ludovica Laschena<sup>1</sup> · Marco Bicchetti<sup>1</sup> · Emanuele Messina<sup>1</sup> · Sara Lucciola<sup>1</sup> · Carlo Catalano<sup>1</sup> · Valeria Panebianco<sup>1</sup>

Received: 1 June 2022 / Accepted: 30 August 2022 / Published online: 17 September 2022

© The Author(s) 2022

## Abstract

**Objective** To investigate the impact of an artificial intelligence (AI) software and quantitative ADC (qADC) on the inter-reader agreement, diagnostic performance, and reporting times of prostate biparametric MRI (bpMRI) for experienced and inexperienced readers.

**Materials and methods** A total of 170 multiparametric MRI (mpMRI) of patients with suspicion of prostate cancer (PCa) were retrospectively reviewed by one experienced and one inexperienced reader three times, following a wash-out period. First, only the bpMRI sequences, including T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI) sequences, and apparent diffusion coefficient (ADC) maps, were used. Then, bpMRI and quantitative ADC values were used. Lastly, bpMRI and the AI software were used. Inter-reader agreement between the two readers and between each reader and the mpMRI original reports was calculated. Detection rates and reporting times were calculated for each group.

**Results** Inter-reader agreement with respect to mpMRI was moderate for bpMRI, Quantib, and qADC for both the inexperienced (weighted  $k$  of 0.42, 0.45, and 0.41, respectively) and the experienced radiologists (weighted  $k$  of 0.44, 0.46, and 0.42, respectively). Detection rate of PCa was similar between the inexperienced (0.24, 0.26, and 0.23) and the experienced reader (0.26, 0.27 and 0.27), for bpMRI, Quantib, and qADC, respectively. Reporting times were lower for Quantib (8.23, 7.11, and 9.87 min for the inexperienced reader and 5.62, 5.07, and 6.21 min for the experienced reader, for bpMRI, Quantib, and qADC, respectively).

**Conclusions** AI and qADC did not have a significant impact on the diagnostic performance of both readers. The use of Quantib was associated with lower reporting times.

**Keywords** Multiparametric MRI · Biparametric MRI · Prostate cancer · Quantitative ADC · Artificial intelligence · Inter-reader agreement

## Introduction

Prostate multiparametric MRI (mpMRI) is the most accurate imaging study for prostate cancer (PCa) diagnosis, and it is increasingly used worldwide for early detection, staging, follow-up, and active surveillance [1–3]. The PI-RADS recommendations, describing a standardized protocol and

reporting system [4], have contributed to the very high diagnostic performance in PCa detection reported by several level 1 evidence trials and a Cochrane systematic review [5–9]. An alternative to mpMRI is an abbreviated protocol known as biparametric prostate MRI (bpMRI), which combines T2-weighted imaging (T2WI) and diffusion-weighted imaging (DWI)/apparent diffusion coefficient (ADC) without using contrast media [8, 10–12]. bpMRI has the advantages of avoiding the costs and potential side effects related to the use of contrast media, and of shortening the acquisition times. It has been shown that bpMRI has a high accuracy when interpreted by experienced radiologists, whereas the available evidence suggests lower performance for less experienced readers [13–16]. Furthermore, the use of a

✉ Valeria Panebianco  
Valeria.panebianco@uniroma1.it

<sup>1</sup> Department of Radiological Sciences, Oncology and Pathology, Sapienza University/Policlinico Umberto I, Viale Regina Elena 324, 00161 Rome, Italy

shorter bpMRI protocol only partly addresses the problems related to the heavy clinical workload for the genitourinary radiologist. With the growing demand for prostate MRI, it is imperative to maximize the diagnostic potential of bpMRI and at the same time to optimize reporting times for each patient [17–19].

Several possible approaches may be investigated to improve the diagnostic potential of bpMRI for less experienced radiologists. An example could involve the use of quantitative data extracted from the ADC map as a decision support system for the interpretation of equivocal lesions. The PI-RADS guidelines state that diffusion-weighted imaging is the dominant sequence for scoring lesions in the peripheral zone of the prostate, which is the most common site where PCa lesions arise. The calculation of quantitative ADC (qADC) values for a particular lesion could improve the diagnostic confidence of the radiologist in scoring suspicious lesions [20–25]. This is especially the case in bpMRI where the absence of dynamic contrast enhancement (DCE) could make equivocal cases much more difficult to interpret [26].

Another potential approach to enhance the diagnostic performance of bpMRI could be the use of artificial intelligence (AI) as a computer-aided diagnosis (CAD) system. A CAD system can serve as a complete automated diagnosis tool, or as a support tool for the radiologist with the goal of improving diagnostic accuracy and/or productivity [27, 28]. In recent years, numerous machine learning (ML) and deep learning (DL) algorithms have been applied to medical imaging and prostate MRI, mostly in preliminary research settings [29–31]. AI algorithms can be adapted to a variety of different tasks in prostate imaging, including quality control, segmentation, detection, and characterization [27].

The aim of this study was to evaluate the use of qADC measurements and an AI-based CE-approved software (Quantib Prostate) in the interpretation of prostate bpMRI, with focus on inter-reader agreement, performance in detecting PCa, and reporting time.

## Material and methods

### Patient population and MRI protocol

The study retrospectively included mpMRI studies performed at our institution for suspicion of PCa during the period of May 2021 through November 2021, with waiver of informed consent approved by the institutional review board. Inclusion criteria were the availability of all three mpMRI sequences (T2WI, DWI/ADC, DCE) and of the official PI-RADS v2.1-compliant mpMRI report provided by a senior genitourinary radiologist (VP) with 15 years of experience at a high-volume referral center in prostate diagnostics (> 1000 prostate mpMRI studies read per year). Exclusion criteria were represented by inadequate image quality of one of the bpMRI sequences (T2WI, DWI or ADC), assessed according to PI-QUAL parameters [32].

MRI examinations were performed on a 3.0 Tesla MRI (GE Discovery 750, GE Healthcare, Milwaukee, USA), using a 32-channel surface phased-array body coil (TOR-SOPA), with a PI-RADS v2.1-compliant protocol, consisting of a high-resolution T2WI in axial and coronal planes. During DWI, b values were set at 50, 800, and 1500; the ADC map was computed with b values of 50 and 800. Perfusion imaging (DCE) was performed following intravenously by gadobutrol (0.1 mmol/kg). Table 1 contains a detailed list of MRI protocol parameters. The patients were instructed to perform a rectal enema 2–4 h before the test.

### Image interpretation and analysis

MRI studies were independently interpreted by one inexperienced and one experienced radiologist (AF, SC). The inexperienced radiologist had 3 months of experience in prostate imaging, had received fellowship-level training comprising theoretical lectures and practical training including supervised reading of approximately 100 cases, and had

**Table 1** Multiparametric MRI acquisition parameters

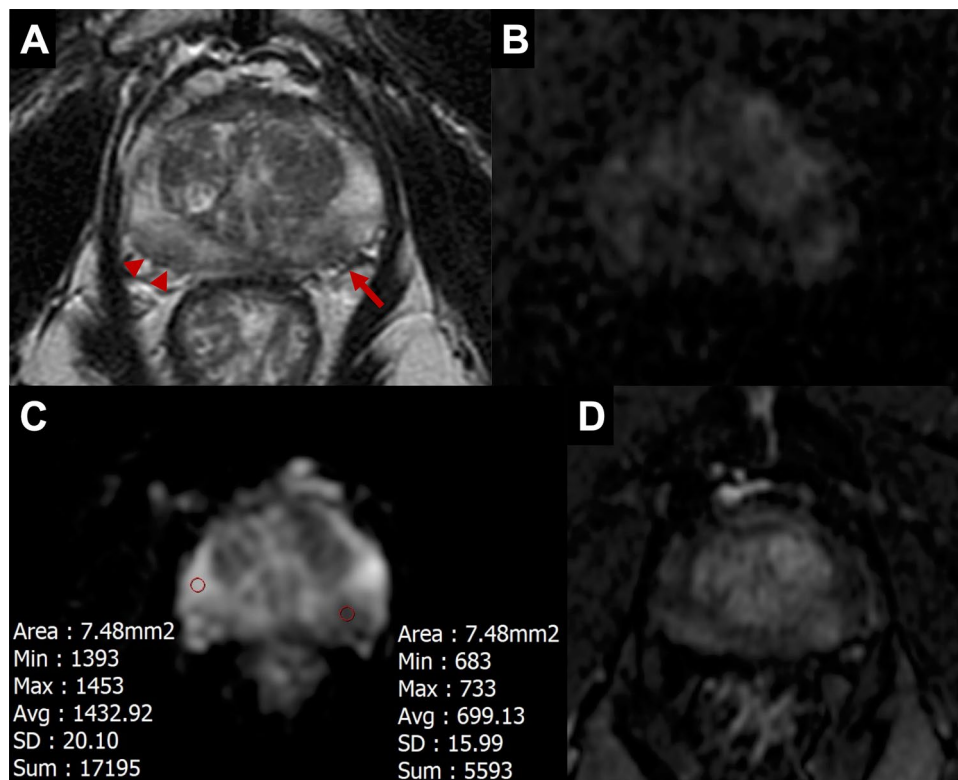
	T2WI	DWI	DCE
Sequence type	Fast Recovery Fast Spin Echo (FRFSE)	Echo Planar Imaging (EPI)	LAVA Gradient-Echo
Acquisition plane	Axial & Coronal	Axial	Axial
Number of averages	6	2 (b 50); 6 (b 800); 12 (b 1500)	1
Slice thickness (mm)	3	3	4
Matrix size	320×224	90×90	160×140
Field of View (cm)	18×18	20×20	18×18
b-values	N/A	50–800–1500	N/A
Temporal resolution (s)	N/A	N/A	6
Contrast media	N/A	N/A	Gadobutrol 0.1 mmol/Kg

independently read 50 mpMRI studies before the beginning of the study. The experienced reader had 5 years of experience in prostate imaging, had completed fellowship training, and had read > 1000 total mpMRI studies (> 200 per year). Both readers, blinded to the mpMRI results, interpreted the bpMRI studies and scored any identified lesions according to PI-RADS v2.1 recommendations. Given that only the T2WI and DWI/ADC sequences were evaluated, lesions in the peripheral zone were scored as if the DCE was negative. Lesion number, location, and PI-RADS score were recorded. Each radiologist interpreted the MRIs three separate times, using a wash-out period of at least three weeks between each reading for memory extinction. The first reading was based on interpretation of the bpMRI sequences (T2WI, DWI, ADC). The second reading was performed using the bpMRI sequences, as well as the normalized ADC value for the suspicious foci to provide additional insight into the suspicion level of each focus. For the calculation of qADC, a circular region of interest (ROI) was placed independently by each radiologist interpreting the study on each suspicious focus in the ADC map at the area corresponding to the highest

signal intensity at high b values in DWI. An equally sized ROI was placed on the normal-appearing peripheral zone or transition zone (according to the location of the suspicious focus) on the same slice as the lesion to normalize the qADC value (Fig. 1). The average pixel value of the lesion ROI was divided by the average pixel value of the normal prostate ROI to yield an ADC ratio. As an additional parameter for scoring suspicious lesions, the resulting ratio was considered as significant for upgrading PI-RADS 3 lesions to PI-RADS 4, using a threshold of 0.59 [26]. The third reading was performed using the Quantib Prostate v1.2.0 software, using the bpMRI sequences, as detailed below. Reporting times were recorded for both radiologists.

### Quantib prostate work-up

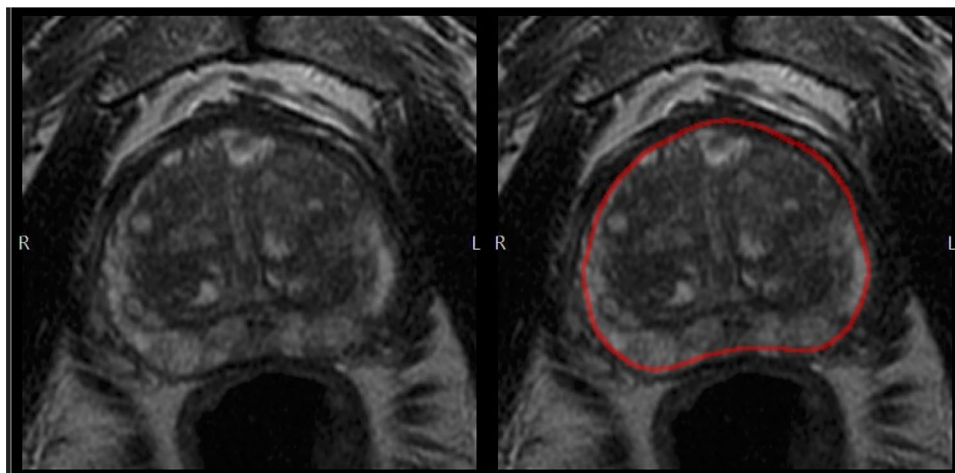
Quantib® Prostate (Quantib BV, Rotterdam, The Netherlands) is an FDA and CE-approved MRI viewing and reporting platform based on deep learning (DL). During the Quantib Prostate work-up, the first step was the semi-automated generation of a segmentation contour of the



**Fig. 1** Sixty-three-year-old biopsy-naïve male with a serum PSA of 5.03 ng/ml and a PSA density of 0.15. A. T2WI showing hypointense foci in the left (arrow) and right (arrowheads) postero-lateral peripheral zone, more apparent on the left. B. DWI at b value of 1500 showing moderate hyperintensity of the two foci, more evident on the left. C. ADC map showing marked hypointensity of the left postero-lateral focus; quantitative ADC measurements reveal an ADC

ratio of 0.49. D. Early phase DCE image revealing enhancement of both foci mpMRI assigned a PI-RADS score of 4 to both foci, based on the positive DCE. On bpMRI with qADC, a PI-RADS score of 4 was assigned to the left focus based on an ADC ratio < 0.59, and a PI-RADS score of 3 was assigned to the right focus, based on an ADC ratio > 0.59. Histopathology showed a Gleason 3+4 PCa on the left side focus and inflammatory changes on the right-side focus.

**Fig. 2** Automatically generated prostate contour overlaid on the axial T2WI is the first step of the Quantib Prostate workflow. The user can manually modify the segmentation, if needed, before approving it



prostate gland based on the T2WI axial sequence (Fig. 2), which was reviewed by the radiologists and manually edited, if necessary. The second step was image interpretation on the DICOM viewing interface of the software that showed both the bpMRI sequences and an automatically generated colorimetric map based on Convolutional Neural Networks (CNNs), overlaid on T2WI images, that shows in different colors the voxels that are more likely to have clinically significant PCa (csPCA, ISUP Grade  $\geq 2$ ) (Fig. 3). In this step, the radiologists were able to identify and score suspicious lesions by directly clicking on them on the MRI images. In the last step, the final report was automatically generated, manually edited if required, and exported by the software.

### Statistical analysis

The inter-reader agreement on lesion score was determined between the two readers for all three interpretation methods as well as between each reader and the mpMRI findings using the weighted Cohen's kappa. In addition, a Cohen's kappa score was also calculated at both patient and lesion levels for presence or absence of suspicious foci (PI-RADS  $\leq 2$  vs. PI-RADS  $\geq 3$ ). Agreement was considered slight for kappa values of 0.00–0.20, fair for values of 0.21–0.40, moderate for values of 0.41–0.60, substantial for values of 0.61–0.80, almost perfect for values of 0.81–1.00. Detection rate was defined as the ratio of detected PCa cases to the total number of cases, using histopathology results of MRI-TRUS fusion targeted biopsy as the reference standard. Detection rate was calculated both in terms of csPCA detection (ISUP Grade  $\geq 2$ ) and overall PCa detection for the three interpretation methods (bpMRI, Quantib, qADC) and for mpMRI. One-way ANOVA was used to compare reporting times among the three patient groups for both the inexperienced and experienced readers. Statistical significance was defined as a *p* value less than 0.05. All statistical analyses

were performed using the R statistical software version 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).

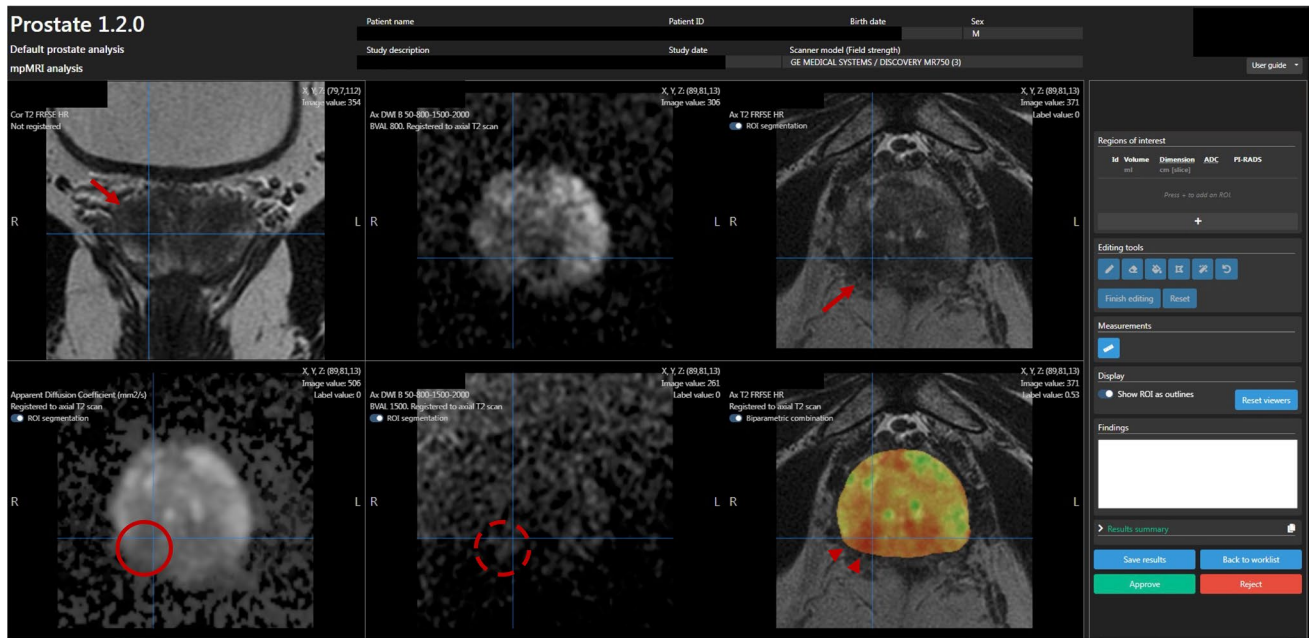
### Results

A total of 204 mpMRI studies were evaluated for inclusion in the study. Thirty-four MRI examinations were excluded due to inadequate quality on at least one of the bpMRI sequences. The 170 patients included in the study were homogeneous between the bpMRI, Quantib, and qADC groups with regard to median age (65 vs. 68 vs. 64 years old,  $p=0.10$ ), median total PSA value (7.6 ng/ml vs. 7.9 ng/ml vs. 7.3 ng/ml,  $p=0.72$ ), and median PSA density (0.13 ng/ml/ml vs. 0.11 ng/ml/ml vs. 0.13 ng/ml/ml,  $p=0.64$ ).

At mpMRI, a total of 238 lesions were identified by the senior radiologist (median of 1 lesion per patient), of which 64 (26.9%) PI-RADS 2 lesions, 73 (30.7%) PI-RADS 3 lesions, 90 (37.8%) PI-RADS 4 lesions, and 12 (5.0%) PI-RADS 5 lesions. Overall, 128/170 patients (75.3%) had at least one lesion classified as PI-RADS 3 or greater. The three study groups had a similar distribution with regards to the PI-RADS scoring of the lesions ( $p=0.99$ , Table 2). One hundred and two patients underwent MRI-TRUS fusion targeted biopsy, of which 35/102 (34.3%) patients had negative results, 17/102 (16.7%) patients had Gleason Score 3 + 3 PCa, and 50/102 (49.0%) patients had Gleason Score  $\geq 3 + 4$  PCa. There were no statistically significant differences among the study groups in the proportion of overall PCa and clinically significant PCa ( $p=0.91$ , Table 2).

The mean qADC value of identified lesions was 0.56 ( $\pm 0.16$  SD) for the inexperienced reader and 0.58 ( $\pm 0.18$  SD) for the experienced reader.

The clinical, radiologic, and pathologic characteristics of the three different patient groups are summarized in Table 2.



**Fig. 3** Seventy-five-year-old biopsy-naïve male with a serum PSA of 12.7 ng/ml and a PSA density of 0.48. Quantib Prostate interface for the viewing/reporting step. The bpMRI sequences are shown, in order: coronal T2WI, DWI at b value of 800, axial T2WI, ADC map, DWI at b value of 1500, colorimetric map overlaid on the axial T2WI. bpMRI images show a hypointense focus on the T2WI, in the right postero-lateral peripheral zone at the prostate base (arrows). The lesion doesn't show hyperintensity on the high b value DWI (dashed

circle). The ADC map shows only moderate hypointensity (solid circle). A PI-RADS score of 3 would have been assigned to the lesion based on the bpMRI sequences. The AI-generated colorimetric map on the bottom right shows a markedly “hot” (red) area corresponding to the right postero-lateral focus (arrowheads), suggesting a high probability focus; the final PI-RADS score assigned was 4, based on the colorimetric map. Histopathology revealed a Gleason score 3+4 PCa

### Inter-reader agreement between experienced and inexperienced readers

The inter-reader agreement for lesion scoring between the experienced and inexperienced readers was fair ( $k=0.38$ ,  $p<0.00001$ ) for bpMRI, moderate ( $k=0.41$ ,  $p<0.00001$ ) for Quantib and moderate ( $k=0.41$ ,  $p<0.00001$ ) for qADC.

The inter-reader agreement for lesion significance between the two readers was moderate ( $k=0.40$ ,  $p<0.00001$ ) for bpMRI, moderate ( $k=0.41$ ,  $p<0.00001$ ) for Quantib, and fair ( $k=0.39$ ,  $p<0.00001$ ) for qADC.

The inter-reader agreement for patient significance between the two readers was moderate for both bpMRI ( $0.42$ ,  $p<0.00001$ ), Quantib ( $0.44$ ,  $p<0.00001$ ), and qADC ( $0.41$ ,  $p<0.00001$ ).

### Inexperienced radiologist analysis

With respect to mpMRI, the inter-reader agreement for the inexperienced radiologist was moderate at both the per-lesion analysis ( $k=0.42$ ,  $p<0.00001$  for bpMRI,  $k=0.45$ ,  $p<0.00001$  for Quantib,  $k=0.41$ ,  $p<0.00001$  for qADC), the significant lesion analysis ( $k=0.44$ ,  $p<0.00001$  for bpMRI,  $k=0.46$ ,  $p<0.00001$  for Quantib,  $k=0.42$ ,

$p<0.00001$  for qADC) and the per-patient analysis ( $k=0.43$ ,  $p<0.00001$  for bpMRI,  $k=0.44$ ,  $p<0.00001$  for Quantib,  $k=0.42$ ,  $p<0.00001$  for qADC), Table 3. Detection rate of PCa was 0.29 for mpMRI, 0.24 for bpMRI, 0.26 for Quantib, 0.23 for qADC. Detection rate of csPCa was 0.20 for mpMRI, 0.16 for bpMRI, 0.17 for Quantib, 0.14 for qADC (Table 4).

The average time needed to the inexperienced radiologist for the overall reporting of MRI examinations was 8,23 min (IQR: 5,32–10,13 min) for the bpMRI, 7,11 min (IQR: 4,43–9,36 min) for Quantib, 9,87 min (IQR: 5,72–12,01 min) for qADC. The difference between the three groups was statistically significant ( $p<0.00001$ ).

### Experienced radiologist analysis

The inter-reader agreement for the experienced radiologist was moderate at both the per-lesion analysis ( $k=0.44$ ,  $p<0.00001$  for bpMRI,  $k=0.46$ ,  $p<0.00001$  for Quantib,  $k=0.42$ ,  $p<0.00001$  for qADC), the significant lesion analysis ( $k=0.43$ ,  $p<0.00001$  for bpMRI,  $k=0.45$ ,  $p<0.00001$  for Quantib,  $k=0.42$ ,  $p<0.00001$  for qADC) and the per-patient analysis ( $k=0.45$ ,  $p<0.00001$  for bpMRI,  $k=0.47$ ,  $p<0.00001$  for Quantib,  $k=0.43$ ,

**Table 2** Patient characteristics

Variable	bpMRI (n=56)	Quantib (n=57)	qADC (n=57)	p Value
Age, years <i>Median (IQR)</i>	65 (60–71)	68 (60–71)	64 (62–75)	0.10
PSA, ng/ml <i>Median (IQR)</i>	7.6 (4.52 – 8.56)	7.9 (5.02 – 8.87)	7.3 (4.63–9.01)	0.72
PSA density, ng/mL/cc <i>Median (IQR)</i>	0.13 (0.09 – 0.16)	0.11 (0.08 – 0.18)	0.11 (0.10–0.19)	0.64
<i>PI-RADS v2, n (%)</i> *				
2	24 (26.1)	25 (26.6)	27 (28.7)	0.99
3	30 (32.6)	29 (30.9)	28 (29.8)	
4	34 (37.0)	35 (37.2)	33 (31.9)	
5	4 (4.3)	5 (5.3)	6 (6.4)	
<i>Histopathology (MRI-TRUS fusion targeted biopsy), n (%)</i>				
Negative	11 (33.3)	13 (37.1)	15 (44.1)	0.91
<i>ciPCa</i>	6 (18.2)	7 (20.0)	5 (14.7)	
<i>csPCa</i>	16 (48.5)	15 (42.9)	14 (41.2)	
Variable	bpMRI (n=56)	Quantib (n=57)	qADC (n=57)	p Value
Age, years <i>Median (IQR)</i>	65 (60–71)	68 (60–71)	64 (62–75)	0.10
PSA, ng/ml <i>Median (IQR)</i>	7.6 (4.52 – 8.56)	7.9 (5.02 – 8.87)	7.3 (4.63–9.01)	0.72
PSA density, ng/mL/cc <i>Median (IQR)</i>	0.13 (0.09 – 0.16)	0.11 (0.08 – 0.18)	0.11 (0.10–0.19)	0.64
<i>PI-RADS v2, n (%)</i> *				
2	24 (26.1)	25 (26.6)	27 (28.7)	0.99
3	30 (32.6)	29 (30.9)	28 (29.8)	
4	34 (37.0)	35 (37.2)	33 (31.9)	
5	4 (4.3)	5 (5.3)	6 (6.4)	
<i>Histopathology (MRI-TRUS fusion targeted biopsy), n (%)</i>				
Negative	11 (33.3)	13 (37.1)	15 (44.1)	0.91
<i>ciPCa</i>	6 (18.2)	7 (20.0)	5 (14.7)	
<i>csPCa</i>	16 (48.5)	15 (42.9)	14 (41.2)	

\* According to the mpMRI report

**Table 3** Summary of results—inter-reader agreement between mpMRI and the three studies

	bpMRI	Quantib	qADC
<i>Inexperienced</i>			
Lesion score	0.42 ( $p < .00001$ )	0.45 ( $p < .00001$ )	0.41 ( $p < .00001$ )
Lesion significance	0.44 ( $p < .00001$ )	0.46 ( $p < .00001$ )	0.42 ( $p < .00001$ )
Patient significance	0.43 ( $p < .00001$ )	0.44 ( $p < .00001$ )	0.42 ( $p < .00001$ )
<i>Experienced</i>			
Lesion score	0.44 ( $p < .00001$ )	0.46 ( $p < .00001$ )	0.42 ( $p < .00001$ )
Lesion significance	0.43 ( $p < .00001$ )	0.45 ( $p < .00001$ )	0.42 ( $p < .00001$ )
Patient significance	0.45 ( $p < .00001$ )	0.47 ( $p < .00001$ )	0.43 ( $p < .00001$ )

$p < 0.00001$  for qADC), Table 3. Detection rate of PCa was 0.29 for mpMRI, 0.26 for bpMRI, 0.27 for Quantib, 0.27 for qADC. Detection rate of csPCa was 0.20 for mpMRI, 0.18 for bpMRI, 0.19 for Quantib, 0.16 for qADC (Table 4).

The average time needed to the experienced radiologist for the overall reporting of MRI examinations was 5,62 min (IQR: 3,54–9,13 min) for the bpMRI, 5,07 min (IQR: 3,43–8,76 min) for Quantib, 6,21 min (IQR: 4,32–12,01 min)

for qADC. The difference between the three groups was statistically significant ( $p = 0.00001$ ).

## Discussion

The results of this study showed that the use of the software Quantib Prostate allowed the radiologist to achieve slightly higher inter-reader agreement with mpMRI, compared to just interpreting bpMRI sequences, both in the case

**Table 4** Summary of results—detection rate

	Overall PCa	csPCa
<i>Inexperienced</i>		
bpMRI	0.24	0.16
Quantib	0.26	0.17
qADC	0.23	0.14
<i>Experienced</i>		
bpMRI	0.26	0.18
Quantib	0.27	0.19
qADC	0.27	0.16
mpMRI	0.29	0.20

of experienced and inexperienced readers. The agreement between the experienced and inexperienced readers was comparable, varying slightly between fair and moderate for bpMRI, Quantib, and qADC, with a minor trend toward higher agreement with the use of Quantib. Neither Quantib Prostate nor quantitative ADC measurements could increase detection rate of PCa with reference to mpMRI, for either experienced or inexperienced readers when interpreting bpMRI studies. The agreement of bpMRI with mpMRI was moderate for both bpMRI, Quantib, and qADC, a data that are in line with the available literature. The fact that detection rate is comparable to that of mpMRI suggests that cases of disagreement did not impact detection of PCa foci.

The use of Quantib Prostate was associated with a shorter reporting time, which is potentially valuable in the clinical workflow, due to the increasing demand for prostate MRI examinations, particularly for the inexperienced user due to the longer times needed to report MRI studies.

Although this study showed comparable diagnostic accuracy, several artificial intelligence and deep learning algorithms have been developed to increase or automate the interpretation of prostate MRIs [33–35]. Due to lack of extensive real-world validation, however, there are no definitive data mandating their use in clinical practice to date [36]. Despite interesting results from earliest reports of AI implementations focusing on automated detection and characterization of PCa, currently there is increasing interest for using AI technology to improve quality control and workflow efficiency in radiology [37].

The decrease in reporting times with the use of Quantib Prostate can be attributed to several factors. First, the automated segmentation allows for automated accurate calculation of the PSA density, a task that would otherwise require the radiologist to take three orthogonal measurements of the prostate gland, calculate the prostate volume, and therefore the PSA density. Second, the colorimetric map could allow for easier/quicker identification of suspicious foci, probably more relevant for the inexperienced reader. Third, once the lesions are identified, the user can use a dedicated tool and click on the

lesion to start an automated segmentation of the lesion. After assigning a location and a PI-RADS score to the lesion(s), a structured report is automatically generated and exported. Furthermore, the colorimetric map calculated by the software could reveal suspicious foci that were not initially evident to the radiology on the bpMRI sequences, as it was noticed to happen occasionally in this study.

The AI software, however, was found to be sensitive to the quality of the overall image and the presence of artifacts in our dataset. Optimal image quality and typical prostate shape are needed for successful prostate segmentation algorithms. Well-defined margins are also necessary for accurate results. A user can manually correct a segmented contour if the segmentation is incorrect, but this will add to the reporting time. Further, the image analysis algorithm also requires adequate image quality to produce a colorimetric map that is accurate and contrasted sufficiently to show suspicious foci.

There are several limitations to this study, including the retrospective single-center nature of this study. Secondly, the PI-RADS scoring system was used to classify lesions using bpMRI, even though the score was originally designed to make full use of the entire mpMRI protocol. Therefore, a full agreement between bpMRI and mpMRI may not be feasible in all cases. Furthermore, the diagnostic performance was assessed in terms of detection rate for those lesions that underwent targeted biopsy. Consequently, it is not possible to estimate the true diagnostic accuracy of all readers, since we may have missed lesions at mpMRI that never underwent targeted biopsy. A true assessment of the diagnostic accuracy would require all patients to have a histology report and/or a relatively long period of clinical and radiological follow-up. In addition, it is highly probable that the performance of the inexperienced reader likely improved during the data collection stage; therefore, the calculated inter-reader agreement and detection rate represent an average representative value along the radiologist's learning curve.

Lastly, all prostate MRI scans used in the study were obtained from a high-volume referral center with extensive experience in prostate MRI and an optimized acquisition protocol. Considering that the AI-based algorithms are quite sensitive to artifacts and degradation of image quality, the findings of this study might not be generalized to different clinical settings.

Future studies should determine whether Quantib Prostate can facilitate a faster learning curve for radiologists with limited experience in genitourinary radiology. In addition, the impact of the use of Quantib Prostate could be investigated for mpMRI in upcoming studies.

## Conclusions

In conclusion, in both experienced and inexperienced readers, the deep learning-based software Quantib Prostate was associated with slightly higher inter-reader agreement with mpMRI. Both the use of Quantib and quantitative ADC achieved similar diagnostic performance in terms of detection rate compared to using only bpMRI sequences. When using Quantib Prostate, both experienced and inexperienced readers could report bpMRI scans in a shorter amount of time.

**Funding** Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. The authors declare that no funds, grants, or other support was received during the preparation of this manuscript.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose. The authors declare that they have no conflict of interest.

**Informed consent** This article does not contain any studies with human participants or animals performed by any of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Mottet N, van den Bergh RCN, Briers E et al (2021) EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer—2020 Update. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol* 79(2):243–262. <https://doi.org/10.1016/j.eururo.2020.09.042>
- US Preventive Services Task Force, Grossman DC, Curry SJ et al (2018) Screening for prostate cancer: US preventive services task force recommendation statement. *JAMA*. 319(18):1901. <https://doi.org/10.1001/jama.2018.3710>
- Panebianco V, Valerio MC, Giuliani A et al (2018) Clinical utility of multiparametric magnetic resonance imaging as the first-line tool for men with high clinical suspicion of prostate cancer. *Eur Urol Oncol* 1(3):208–214. <https://doi.org/10.1016/j.euo.2018.03.008>
- Turkbey B, Rosenkrantz AB, Haider MA et al (2019) Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *Eur Urol* 76(3):340–351. <https://doi.org/10.1016/j.eururo.2019.02.033>
- Drost FH, Osses DF, Nieboer D et al (2019) Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. *Cochrane database of systematic reviews*. John Wiley & Sons, US
- Rouvière O, Puech P, Renard-Penna R et al (2019) Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naïve patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. *Lancet Oncol* 20(1):100–109. [https://doi.org/10.1016/S1470-2045\(18\)30569-2](https://doi.org/10.1016/S1470-2045(18)30569-2)
- Kasivisvanathan V, Rannikko AS, Borghi M et al (2018) MRI-targeted or standard biopsy for prostate-cancer diagnosis. *N Engl J Med* 378(19):1767–1777. <https://doi.org/10.1056/NEJMoa1801993>
- Rais-Bahrami S, Siddiqui MM, Vourganti S et al (2015) Diagnostic value of biparametric magnetic resonance imaging (MRI) as an adjunct to prostate-specific antigen (PSA)-based detection of prostate cancer in men without prior biopsies: Biparametric MRI for prostate cancer detection. *BJU Int* 115(3):381–388. <https://doi.org/10.1111/bju.12639>
- 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. *The Lancet* 389(10071):815–822. [https://doi.org/10.1016/S0140-6736\(16\)32401-1](https://doi.org/10.1016/S0140-6736(16)32401-1)
- van der Leest M, Cornel E, Israël B et al (2019) Head-to-head comparison of transrectal ultrasound-guided prostate biopsy versus multiparametric prostate resonance imaging with subsequent magnetic resonance-guided biopsy in biopsy-naïve men with elevated prostate-specific antigen: a large prospective multicenter clinical study. *Eur Urol* 75(4):570–578. <https://doi.org/10.1016/j.eururo.2018.11.023>
- Woo S, Suh CH, Kim SY, Cho JY, Kim SH, Moon MH (2018) Head-to-head comparison between biparametric and multiparametric MRI for the diagnosis of prostate cancer: a systematic review and meta-analysis. *Am J Roentgenol* 211(5):W226–W241. <https://doi.org/10.2214/AJR.18.19880>
- Bass EJ, Pantovic A, Connor M et al (2021) A systematic review and meta-analysis of the diagnostic accuracy of biparametric prostate MRI for prostate cancer in men at risk. *Prostate Cancer Prostatic Dis* 24(3):596–611. <https://doi.org/10.1038/s41391-020-00298-w>
- Di Campli E, Delli Pizzi A, Seccia B et al (2018) Diagnostic accuracy of biparametric vs multiparametric MRI in clinically significant prostate cancer: Comparison between readers with different experience. *Eur J Radiol* 101:17–23. <https://doi.org/10.1016/j.ejrad.2018.01.028>
- Choi MH, Kim CK, Lee YJ, Jung SE (2019) Prebiopsy biparametric mri for clinically significant prostate cancer detection with PI-RADS version 2: a multicenter study. *Am J Roentgenol* 212(4):839–846. <https://doi.org/10.2214/AJR.18.20498>
- Thestrup KCD, Logager V, Baslev I, Møller JM, Hansen RH, Thomsen HS (2016) Biparametric versus multiparametric MRI in the diagnosis of prostate cancer. *Acta Radiologica Open* 5(8):205846011666304. <https://doi.org/10.1177/2058460116663046>
- Gatti M, Faletti R, Callaris G et al (2019) Prostate cancer detection with biparametric magnetic resonance imaging (bpMRI) by readers with different experience: performance and comparison with multiparametric (mpMRI). *Abdom Radiol* 44(5):1883–1893. <https://doi.org/10.1007/s00261-019-01934-3>
- Caglic I, Sushentsev N, Shah N, Warren AY, Lamb BW, Barrett T (2021) Comparison of biparametric versus multiparametric prostate MRI for the detection of extracapsular extension and



- seminal vesicle invasion in biopsy naïve patients. *Eur J Radiol* 141:109804. <https://doi.org/10.1016/j.ejrad.2021.109804>
18. Pesapane F, Acquasanta M, Meo R et al (2021) Comparison of sensitivity and specificity of biparametric versus multiparametric prostate mri in the detection of prostate cancer in 431 men with elevated prostate-specific antigen levels. *Diagnostics* 11(7):1223. <https://doi.org/10.3390/diagnostics11071223>
  19. Stanzione A, Ponsiglione A, Cuocolo R et al (2019) Abbreviated protocols *versus* multiparametric MRI for assessment of extraprostatic extension in prostatic carcinoma: a multireader study. *Anticancer Res* 39(8):4449–4454. <https://doi.org/10.21873/anticancer.13617>
  20. Gaur S, Harmon S, Rosenblum L et al (2018) Can apparent diffusion coefficient values assist PI-RADS version 2 DWI scoring? A correlation study using the PI-RADSV2 and international society of urological pathology systems. *Am J Roentgenol* 211(1):W33–W41. <https://doi.org/10.2214/AJR.17.18702>
  21. Nguyen TB, Ushinsky A, Yang A et al (2018) Utility of quantitative apparent diffusion coefficient measurements and normalized apparent diffusion coefficient ratios in the diagnosis of clinically significant peripheral zone prostate cancer. *BJR*. <https://doi.org/10.1259/bjr.20180091>
  22. Barrett T, Priest AN, Lawrence EM et al (2015) Ratio of tumor to normal prostate tissue apparent diffusion coefficient as a method for quantifying DWI of the prostate. *Am J Roentgenol* 205(6):W585–W593. <https://doi.org/10.2214/AJR.15.14338>
  23. Bajgirani AM, Mirak SA, Sung K, Sisk AE, Reiter RE, Raman SS (2019) Apparent diffusion coefficient (ADC) Ratio versus conventional ADC for detecting clinically significant prostate cancer With 3-T MRI. *Am J Roentgenol* 213(3):W134–W142. <https://doi.org/10.2214/AJR.19.21365>
  24. De Cobelli F, Ravelli S, Esposito A et al (2015) Apparent diffusion coefficient value and ratio as noninvasive potential biomarkers to predict prostate cancer grading: comparison with prostate biopsy and radical prostatectomy specimen. *Am J Roentgenol* 204(3):550–557. <https://doi.org/10.2214/AJR.14.13146>
  25. Alessandrino F, Taghipour M, Hassanzadeh E et al (2019) Predictive role of PI-RADSV2 and ADC parameters in differentiating Gleason pattern 3 + 4 and 4 + 3 prostate cancer. *Abdom Radiol* 44(1):279–285. <https://doi.org/10.1007/s00261-018-1718-6>
  26. Brancato V, Di Costanzo G, Basso L et al (2020) Assessment of DCE utility for PCa diagnosis using PI-RADS v2 1: effects on diagnostic accuracy and reproducibility. *Diagnostics*. 10(3):164. <https://doi.org/10.3390/diagnostics10030164>
  27. Cuocolo R, Cipullo MB, Stanzione A et al (2019) Machine learning applications in prostate cancer magnetic resonance imaging. *Eur Radiol Exp* 3(1):35. <https://doi.org/10.1186/s41747-019-0109-2>
  28. Bardis MD, Houshyar R, Chang PD et al (2020) Applications of artificial intelligence to prostate Multiparametric MRI (mpMRI): current and emerging trends. *Cancers* 12(5):1204. <https://doi.org/10.3390/cancers12051204>
  29. Gassenmaier S, Afat S, Nickel D, Mostapha M, Herrmann J, Othman AE (2021) Deep learning–accelerated T2-weighted imaging of the prostate: Reduction of acquisition time and improvement of image quality. *Eur J Radiol* 137:109600. <https://doi.org/10.1016/j.ejrad.2021.109600>
  30. Schelb P, Kohl S, Radtke JP et al (2019) Classification of cancer at prostate MRI: deep learning versus clinical PI-RADS assessment. *Radiology* 293(3):607–617. <https://doi.org/10.1148/radiol.2019190938>
  31. Sollini M, Antunovic L, Chiti A, Kirienko M (2019) Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging* 46(13):2656–2672. <https://doi.org/10.1007/s00259-019-04372-x>
  32. Giganti F, Allen C, Emberton M, Moore CM, Kasivisvanathan V (2020) Prostate imaging quality (PI-QUAL): a new quality control scoring system for multiparametric magnetic resonance imaging of the prostate from the PRECISION trial. *Eur Urol Oncol* 3(5):615–619. <https://doi.org/10.1016/j.euo.2020.06.007>
  33. Baltzer PAT, Clauser P (2021) Applications of artificial intelligence in prostate cancer imaging. *Curr Opin Urol* 31(4):416–423. <https://doi.org/10.1097/MOU.0000000000000887>
  34. Turkbey B, Haider MA (2021) Artificial intelligence (AI) for automated cancer detection on prostate MRI opportunities and ongoing challenges from the AJR special series on AI applications. *Am J Roentgenol*. <https://doi.org/10.2214/AJR.21.26917>
  35. Castillo TJM, Arif M, Niessen WJ, Schoots IG, Veenland JF (2020) Automated classification of significant prostate cancer on MRI: a systematic review on the performance of machine learning applications. *Cancers* 12(6):1606. <https://doi.org/10.3390/cancers12061606>
  36. Turkbey B, Haider MA (2022) Deep learning-based artificial intelligence applications in prostate MRI: brief summary. *BJR* 95(1131):20210563. <https://doi.org/10.1259/bjr.20210563>
  37. Gassenmaier S, Küstner T, Nickel D et al (2021) Deep learning applications in magnetic resonance imaging: has the future become present? *Diagnostics* 11(12):2181. <https://doi.org/10.3390/diagnostics11122181>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.