

Mdprostate AI and radiologist performance and interobserver agreement using bpMRI to predict the likelihood of prostate cancer based on PI-RADS

Eduardo Valencia-Segura^{ID}, Beatriz Elias-Perez^{*ID}, Arturo Hernandez-Medina^{ID}
and Francisco J. Omaña-Villagran^{ID}

Department of Magnetic Resonance Imaging, Hospital Angeles Lomas, Huixquilucan, State of Mexico, Mexico

ABSTRACT

Introduction: Artificial intelligence (AI) has improved the diagnostic performance of prostate magnetic resonance imaging (MRI) and reduced interobserver variability in Prostate Imaging Reporting and Data System (PI-RADS) scores. This study compared the diagnostic performance of a radiologist and mdprostate AI, using biparametric (bp) MRI, for predicting the likelihood of clinically significant prostate cancer (csPCa) based on PI-RADS scores and assessed interobserver agreement.

Materials and methods: This retrospective cross-sectional study included patients who underwent bpMRI for suspected prostate cancer. bpMRI scans were analyzed using PI-RADS v2.1 by an experienced radiologist and mdprostate. Prostate lesions were classified as low (PI-RADS 1 to 3) or high/very high (PI-RADS 4 to 5) likelihood of csPCa. Sensitivity, specificity, positive and negative predictive values, likelihood ratios, and accuracy were calculated. Cohen's kappa coefficient was used to assess interobserver agreement, and Bowker's test of symmetry was used to analyze systematic differences in ordinal PI-RADS scoring. **Results:** Eighty-two men with a mean age of 64 ± 9.9 years (range 43-85) were included. The mean PSA was 9.9 ± 10.3 ng/mL. Mdprostate showed a sensitivity of 96.3%, specificity of 89.1%, and overall accuracy of 91.5% in predicting high and very high likelihood of csPCa, with the radiologist as the gold standard. Interobserver agreement between the radiologist and mdprostate was almost perfect ($k = 0.81$; 95% CI, 0.67-0.96). Bowker's test showed significant differences in PI-RADS categories ($p = 0.017$), indicating that mdprostate tended to assign higher PI-RADS scores than the radiologist, particularly for PI-RADS 5 lesions. **Conclusion:** Mdprostate demonstrated high diagnostic performance compared to the radiologist and near-perfect agreement between both, using bpMRI to predict a high or very high likelihood of csPCa based on PI-RADS assessment.

Keywords: Prostate cancer. Artificial intelligence. Biparametric magnetic resonance imaging. Mdprostate. Prostate Imaging Reporting and Data System.

INTRODUCTION

Prostate cancer is one of the most common male malignancies worldwide, and an important cause of morbidity and mortality, ranking fourth globally with 1.4 million cases¹. Prostate biparametric magnetic resonance imaging (bpMRI), based on T2-weighted (T2W),

diffusion-weighted imaging (DWI), and apparent diffusion coefficient (ADC) sequences, is faster and more cost-effective than multiparametric (mp) MRI, with comparable diagnostic performance for detecting prostate cancer^{2,3}. Prostate MRI may exhibit high interobserver variability, even among experienced radiologists⁴. In this context, artificial intelligence (AI) is a valuable tool for detecting

*Corresponding author:

Beatriz Elias-Perez
E-mail: betyalias@hotmail.com

Received for publication: 12-12-2025

Accepted for publication: 23-03-2026

DOI: 10.24875/JMEXFRI.M26000132

Available online: 08-07-2026

J Mex Fed Radiol Imaging. 2026;5(2):98-106

www.JMeXFRI.com

2696-8444 / © 2026 Federación Mexicana de Radiología e Imagen, A.C. Published by Permanyer. This is an open access article under the CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

and characterizing prostate lesions, offering diagnostic performance comparable to that of expert radiologists with adequate reproducibility^{5,6}.

Commercial platforms designed for clinical integration, such as Quantib Prostate (2020), ProstatID (2022), mdprostate (2023), and QP-Prostate (2024), incorporate automated prostate segmentation, suspicious lesion detection, Prostate Imaging Reporting and Data System (PI-RADS) scoring, and structured reporting. These tools improve diagnostic sensitivity, optimize reading times, and especially assist radiologists with less experience in interpreting prostatic mpMRI⁷⁻¹⁰. Mdprostate (mediaire GmbH, Berlin, Germany) is an AI-based software designed to assist in interpreting prostate MRI using biparametric sequences, especially T2W, DWI, and ADC^{8,11}. Mdprostate uses deep neural networks for automatic prostate segmentation, calculates quantitative parameters such as prostate volume, lesion dimensions, ADC values, and prostate-specific antigen (PSA) density, detects suspicious prostate lesions, classifies PI-RADS v2.1, and generates structured reports with sectorial maps integrated in the PACS.

Despite AI's high performance in bpMRI, its external validity and formal adoption need to be standardized, with potential biases arising from the training datasets used for the algorithms^{9,10}. Several reviews on AI in MRI emphasize the need for multicenter, prospective validations before widespread clinical adoption^{5,12}. This study compared the diagnostic performance of a radiologist and mdprostate AI using bpMRI to predict the likelihood of csPCa based on PI-RADS score and assessed interobserver agreement between the two.

MATERIALS AND METHODS

This cross-sectional study was conducted from January to October 2025 in the Department of Magnetic Resonance Imaging at Hospital Angeles Lomas, Huixquilucan, State of Mexico, Mexico. Adult men with PSA > 3 ng/mL who underwent prostate bpMRI were included. Patients with claustrophobia, incomplete bpMRI scans, studies with motion artifacts, or without a PI-RADS score generated by mdprostate were excluded. All participants provided written informed consent. The study was approved by the Institutional Research Ethics Committee and Research Committee.

Development and study variables

We recorded the patients' age and PSA (ng/mL). The radiologist and mdprostate assigned a PI-RADS v2.1

score¹³ on bpMRI using a five-point scale to indicate the likelihood of csPCa, with a score of 1 indicating a very low likelihood and a score of 5 indicating a very high likelihood. Individual scores were assigned to each sequence, with an overall score that summarized the likelihood of csPCa. In this study, the PI-RADS score was dichotomized as follows: PI-RADS of 1 to 3 indicated a low likelihood of csPCa; PI-RADS ≥ 4 indicated a high or very high likelihood of csPCa. The PI-RADS score agreement between the radiologist (EVS) with 5 years of MRI experience and mdprostate was assessed.

Imaging acquisition and data analysis protocol

bpMRI was performed on 1.5T MAGNETOM Aera or 3.0T MAGNETOM Skyra Fit equipment (Siemens Healthineers, Erlangen, Germany) with an 18-channel radiofrequency surface coil. A conventional T2W sequence was acquired in the axial plane and DWI ($b \geq 1500$ s/mm²) with ADC map generation. Standard clinical acquisition parameters are shown in table 1.

After acquisition, the images were automatically transferred to the mdprostate software (Mediaire GmbH, Berlin, Germany) for analysis. Mdprostate generated a PI-RADS score, and the radiologist's evaluation was recorded.

Statistical analysis

Quantitative variables were analyzed with descriptive statistics. The Shapiro-Wilk test was used to assess the distribution of quantitative variables. Results were expressed as mean and standard deviation (SD). Diagnostic performance using the radiologist or mdprostate as the gold standard was assessed with sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive and negative likelihood ratios, and accuracy. Inter-reader agreement between the radiologist and mdprostate was analyzed using Cohen's kappa, and Bowker's test of symmetry was used to evaluate systematic differences in paired ordinal PI-RADS scores. A p value < 0.05 was considered statistically significant. The analysis was performed with RKWard v0.8.0¹⁴.

RESULTS

Eighty-two patients with a mean age of 64 ± 9.9 years (range 43-85) were included. The mean PSA value was 9.9 ± 10.3 ng/mL, with a median of 6.1 ng/mL (range 3.1-71.9 ng/mL).

Table 1. Prostate bpMRI protocol used by mdprostate AI

Field strength	Sequence	FOV (mm)	Matrix	TR (ms)	TE (ms)	b values [number of samples] (s/mm ²)	Slices/Slice thickness (mm) ^a
1.5T	Axial T2W TSE	220*220	266*320	5730	87	N/A	30/3.0
	Axial EPI DWI	220*220	110*110	6700	74	50 [2], 400 [5], 1600 [10], 2000 ^b	30/3.0
3.0T	Axial T2W TSE	200*200	352*352	6000	105	N/A	25/3.0
	Axial EPI DWI	200*200	104*104	4600	75	50 [2], 800 [5], 1500 [14], 2000 ^b	25/3.0

^aNo gap between slices; ^bSynthetic high b value was calculated by the scanner. DWI: diffusion-weighted imaging; EPI: echo planar imaging; FOV: field-of-view; MRI: magnetic resonance imaging; N/A: not applicable; T: tesla; TE: echo time; TI: inversion-recovery time; TR: repetition time; TSE: turbo spin echo; T2W: T2-weighted.

Comparison of PI-RADS scores assigned by the radiologist and mdprostate

Cases classified as PI-RADS 2 (n = 51, 62.2%) and PI-RADS 3 (n = 6, 7.3%) by the radiologist were comparable to those classified by mdprostate (n = 43, 52.4% and n = 5, 6.1%, respectively) (Table 2). The radiologist classified 14 (17.1%) and 11 (13.4%) cases as PI-RADS 4 and 5, respectively, while mdprostate classified 9 (11.0%) and 21 (25.6%) cases. The radiologist and mdprostate PI-RADS scores were compared using Bowker's test of symmetry with a significant difference (p = 0.017), indicating systematic differences in score assignment. Specifically, mdprostate classified more cases as higher (PI-RADS 5) than the radiologist. This difference is likely due to mdprostate's tendency to overestimate lesion segmentation. However, this overestimation was observed from categories 4 to 5 and does not directly impact the likelihood of csPCa.

Diagnostic performance of mdprostate using bpMRI, with the radiologist as the gold standard in predicting the likelihood of csPCa

Compared with the radiologist, mdprostate showed a sensitivity of 96.3% (95% CI: 92.1-100) and a specificity of 89.1% (95% CI: 82.3-95.8) for predicting high likelihood (PI-RADS 4) and very high likelihood (PI-RADS 5) of csPCa (Table 3). The PPV was 81.2%, and the NPV was 98.0%, reflecting a high probability that both positive and negative classifications were accurate. The diagnostic accuracy was 91.4% (95% CI: 85.4-97.5).

The positive likelihood ratio was 8.8 (4.1-18.9), indicating that a positive result strongly predicts csPCa. The negative likelihood ratio was 0.04 (0.01-0.29), indicating that negative results greatly reduce the probability of disease. Mdprostate tended to classify more cases as very high likelihood of csPCa (PI-RADS 5).

Table 2. Comparison of PI-RADS scores assigned by the radiologist and mdprostate in patients with suspicious prostate lesions

Score	Radiologist (n = 82)	Mdprostate (n = 82)
PI-RADS 1, n (%)	0	4 (4.9)
PI-RADS 2, n (%)	51 (62.2)	43 (52.4)
PI-RADS 3, n (%)	6 (7.3)	5 (6.1)
PI-RADS 4, n (%)	14 (17.1)	9 (11.0)
PI-RADS 5, n (%)	11 (13.4)	21 (25.6)

PI-RADS: Prostate Imaging Reporting and Data System.

Diagnostic performance of the radiologist using bpMRI, with mdprostate as the gold standard in predicting the likelihood of csPCa

The radiologist demonstrated high diagnostic performance compared to mdprostate (Table 4). Sensitivity was 81.2% (95% CI: 63.6-92.8), and specificity was 98.0% (95% CI: 89.3-99.9), indicating an excellent ability of the radiologist to identify true-negative cases. The PPV and NPV were 95.1% and 91.7%, respectively, reflecting a high probability that both positive and negative results were accurate. The accuracy was 92.6% (95% CI 84.7-97.2). The mean positive likelihood ratio was 40.6 (5.8-284.8), suggesting a strong ability to confirm disease, while the mean negative likelihood ratio was 0.19 (0.09-0.39), indicating a good ability to establish a negative diagnosis, though less strong than the positive ratio.

Interobserver agreement in PI-RADS scoring between the radiologist and mdprostate

There was strong concordance between the radiologist and mdprostate, with almost perfect agreement (k = 0.81, 95% CI 0.67-0.96) in assigning PI-RADS scores. The kappa values support the reliability of mdprostate in a

Table 3. Diagnostic performance of mdprostate using bpMRI, with the radiologist as the gold standard in predicting the likelihood of csPCa based on PI-RADS scores

Description	Parameter
Sensitivity, % (95% CI)	96.3 (92.1-100)
Specificity, % (95% CI)	89.1 (82.3-95.8)
PPV, % (95% CI)	81.2 (72.7-89.7)
NPV, % (95% CI)	98.0 (94.9-100)
LR+, mean (min-max)	8.8 (4.1-18.9)
LR-, mean (min-max)	0.04 (0.01-0.29)
Accuracy, % (95% CI)	91.4 (85.4-97.5)

AI: artificial intelligence; bpMRI: biparametric magnetic resonance imaging; csPCa: clinically significant prostate cancer; CI: confidence interval; LR+: positive likelihood ratio; LR-: negative likelihood ratio; NPV: negative predictive value; PI-RADS: Prostate Imaging Reporting and Data System; PPV: positive predictive value.

Table 4. Diagnostic performance of the radiologist using bpMRI, with mdprostate as the gold standard in predicting the likelihood of csPCa based on PI-RADS scores

Description	Parameter
Sensitivity, % (95% CI)	81.2 (63.6-92.8)
Specificity, % (95% CI)	98.0 (89.3-99.9)
PPV, % (95% CI)	95.1 (73.1-99.2)
NPV, % (95% CI)	91.7 (84.4-95.8)
LR+, mean (min-max)	40.6 (5.8-284.8)
LR-, mean (min-max)	0.19 (0.09-0.39)
Accuracy, % (95% CI)	92.6 (84.7-97.2)

AI: artificial intelligence; bpMRI: biparametric magnetic resonance imaging; csPCa: clinically significant prostate cancer; CI: confidence interval; LR+: positive likelihood ratio; LR-: negative likelihood ratio; NPV: negative predictive value; PI-RADS: Prostate Imaging Reporting and Data System; PPV: positive predictive value.

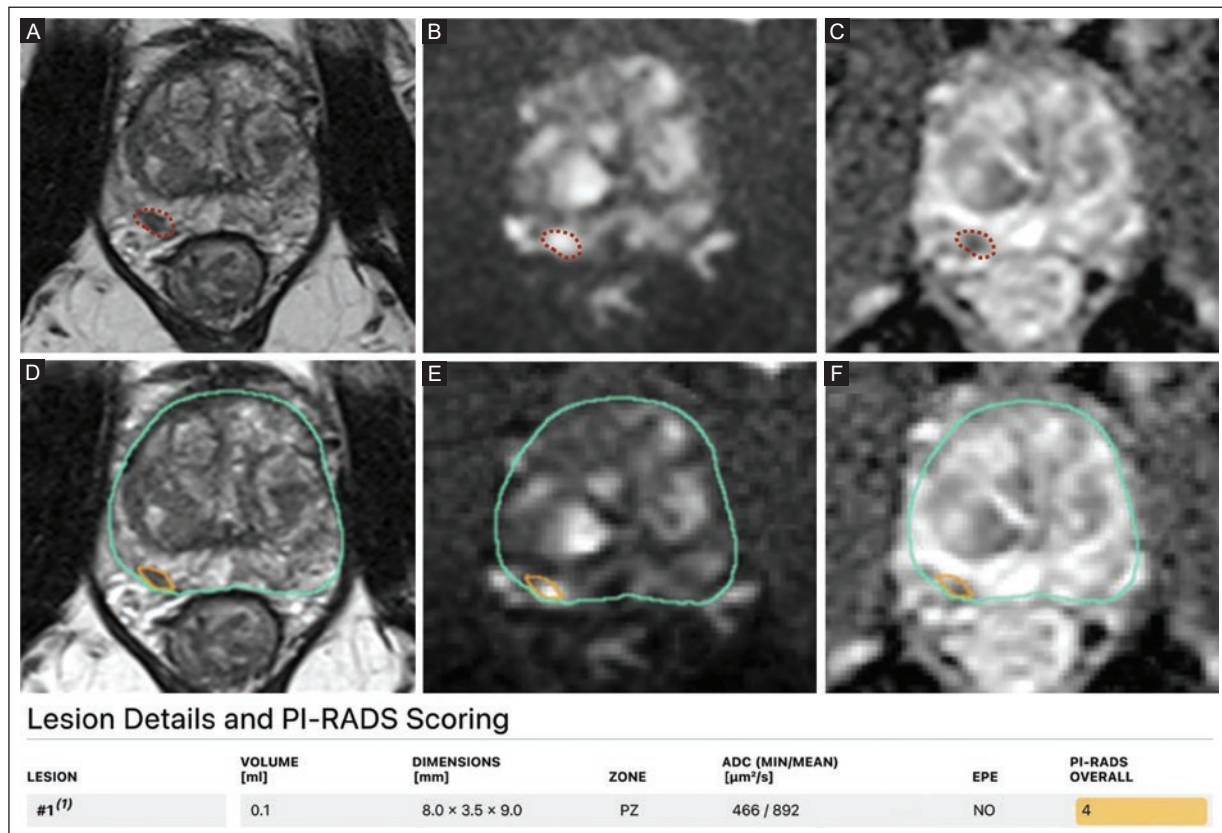


Figure 1. bpMRI of a 69-year-old man with an elevated PSA of 10 ng/mL. The upper panel shows the images assessed by the radiologist: **A:** T2W axial sequence shows a hypointense nodule (dashed line) with well-defined margins, measuring 8.2 mm × 3.6 mm × 6.4 mm along the transverse, anteroposterior, and longitudinal axes, respectively, with an estimated volume of 0.09 mL. The nodule is located in the right posterolateral PZ. **B:** an axial DWI sequence showing marked nodule hyperintensity (dashed line). **C:** ADC map shows marked hypointensity (dashed line) with a mean of 654 $\mu\text{m}^2/\text{s}$ (min 466, max 894), indicating diffusion restriction. The radiologist classified this nodule as PI-RADS 4. The lower panel shows images assessed by mdprostate: **D:** axial T2W sequence, **E:** axial DWI sequence, and **F:** ADC map. Mdprostate analysis shows segmentation of the prostate (green circle) and the nodule (orange circle). The data at the bottom show the prostate nodule volume, dimensions, zone, and ADC, with a PI-RADS score of 4. The histopathologic biopsy diagnosis of the nodule was prostate adenocarcinoma, Gleason 3 + 4. This case showed agreement between the radiologist and mdprostate on PI-RADS scoring.

ADC: apparent diffusion coefficient; bpMRI: biparametric prostate magnetic resonance imaging; DWI: diffusion-weighted imaging; PSA: prostate-specific antigen; PI-RADS: Prostate Imaging Reporting and Data System; PZ: peripheral zone; T2W: T2-weighted.

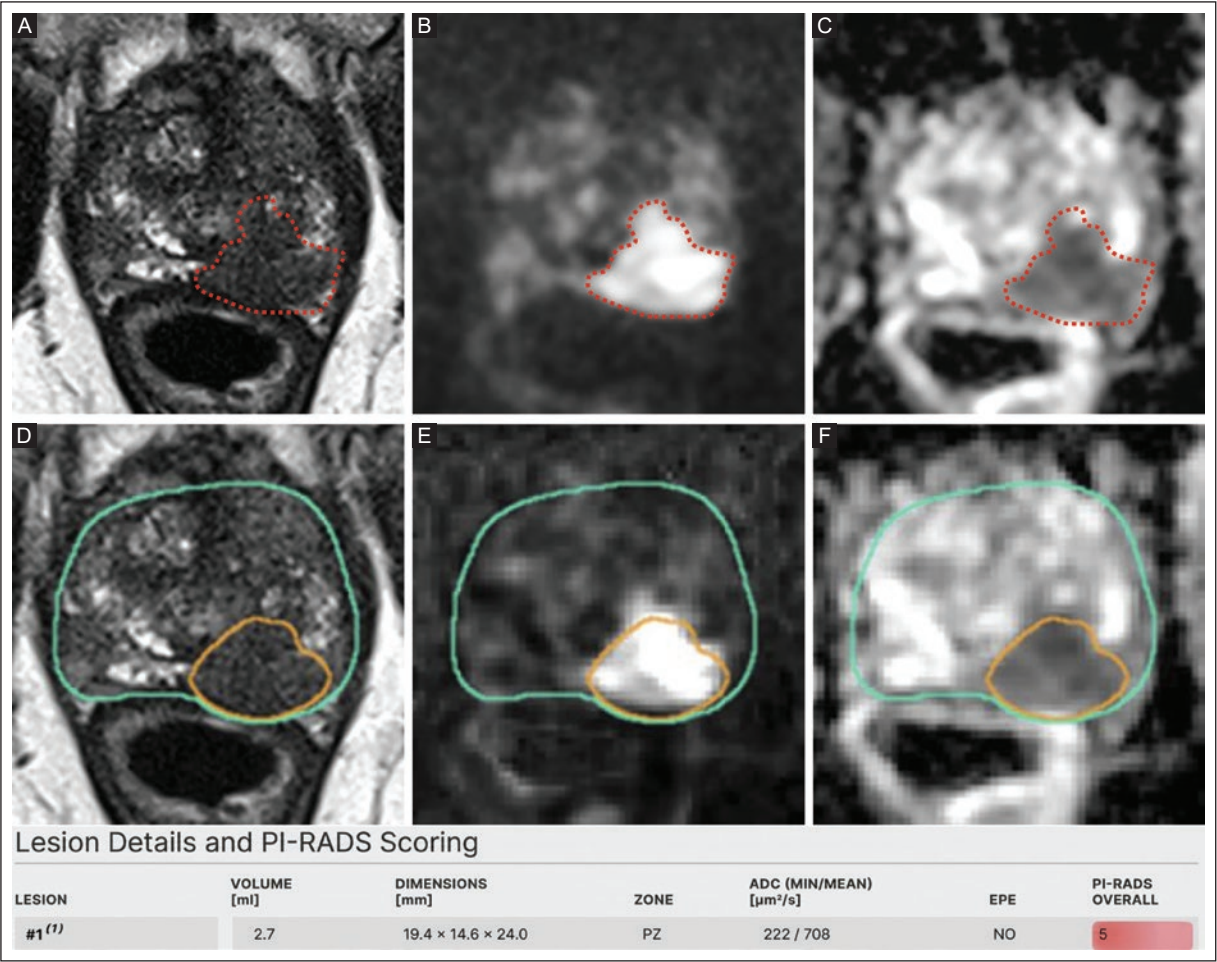


Figure 2. bpMRI of an 85-year-old man with an elevated PSA of 9.5 ng/mL. The upper panel shows the images assessed by the radiologist: **A:** T2W axial sequence shows a nodule with poorly defined margins (dashed line) that partially bulges the prostate capsule, measuring 21.1 mm × 7.2 mm × 16.3 mm along the transverse, anteroposterior, and longitudinal axes, respectively, with an estimated volume of 1.2 mL. The nodule is located in the left posterolateral and posteromedial PZ and extends into the left posterior TZ. **B:** axial DWI sequence shows marked hyperintensity within the nodule (dashed line). **C:** the ADC map shows marked hypointensity (dashed line) due to diffusion restriction, with a mean ADC of 544 $\mu\text{m}^2/\text{s}$ (range 385-750). The radiologist classified the nodule as PI-RADS 5. The lower panel shows the images assessed by mdprostate: **D:** axial T2W sequence, **E:** axial DWI sequence, and **F:** ADC map. Analysis by mdprostate shows segmentation of the prostate (green circle) and nodule (orange circle). The data at the bottom show the volume, dimensions, prostate zone, and ADC of the prostate nodule, with a PI-RADS score of 5. This case showed agreement in PI-RADS scoring between the radiologist and mdprostate.

ADC: apparent diffusion coefficient; bpMRI: biparametric prostate magnetic resonance imaging; DWI: diffusion-weighted imaging; PSA: prostate-specific antigen; PI-RADS: Prostate Imaging Reporting and Data System; PZ: peripheral zone; T2W: T2-weighted; TZ: transition zone.

clinical setting, reinforcing that the likelihood of csPCa was accurately classified in both positive and negative cases when mdprostate was compared with the radiologist.

Figure 1 shows a bpMRI of a 69-year-old man with an elevated PSA of 10 ng/mL. A prostate nodule with marked hyperintensity on the DWI sequence is located in the right posterolateral peripheral zone (PZ). The ADC map shows marked hypointensity, with a mean of 654 $\mu\text{m}^2/\text{s}$, indicating diffusion restriction. The radiologist and mdprostate agreed and classified this nodule as PI-RADS 4. Histopathologic analysis of the nodule biopsy showed prostate adenocarcinoma, Gleason 3 + 4.

Figure 2 shows a bpMRI of an 85-year-old man with an elevated PSA of 9.5 ng/mL. A prostate nodule with marked hyperintensity on the DWI sequence is located in the left posterolateral and posteromedial PZ, extending into the left posterior transition zone (TZ). The ADC map shows marked hypointensity due to diffusion restriction, with a mean ADC of 544 $\mu\text{m}^2/\text{s}$. The radiologist and mdprostate agreed and classified the nodule as PI-RADS 5.

Figure 3 shows a bpMRI of a 62-year-old man with an elevated PSA of 40 ng/mL, fever, and dysuria. The T2W axial sequence shows homogeneous diffuse hypointensity in the TZ and PZ of both prostatic lobes,

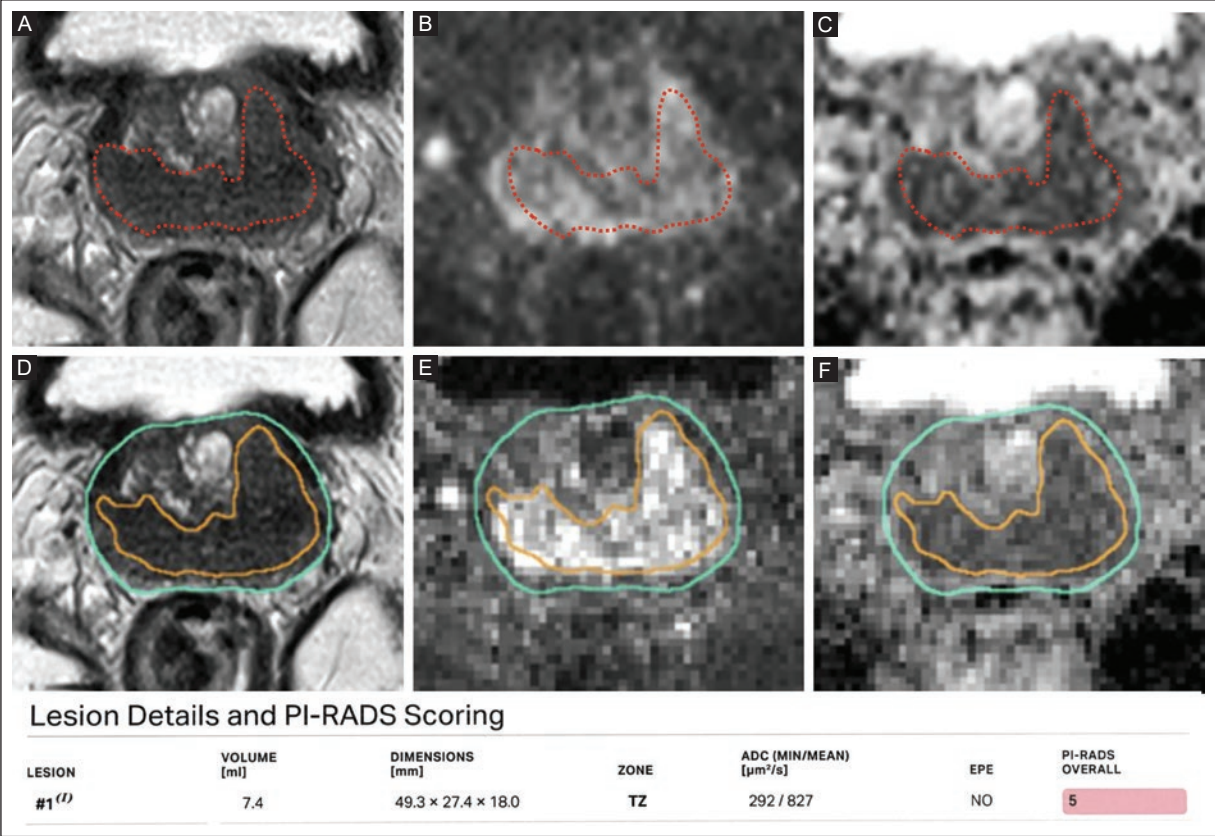


Figure 3. bpMRI of a 62-year-old man with fever, dysuria, and an elevated PSA of 40 ng/mL. The upper panel shows the images assessed by the radiologist: **A:** T2W axial sequence shows homogeneous diffuse hypointensity (dashed line) in the TZ and PZ of both prostate lobes, predominantly on the left, measuring 40 mm × 20.9 mm × 25.5 mm along the transverse, anteroposterior, and longitudinal axes, respectively, with an estimated volume of 11 mL. **B:** axial DWI sequence shows mild TZ and PZ hyperintensity of both prostate lobes, predominantly on the left (dashed line). **C:** ADC map shows marked TZ and PZ hypointensity of both prostatic lobes (dashed line), predominantly on the left, with restricted diffusion and a mean ADC of 762 $\mu\text{m}^2/\text{s}$ (range 520-984). The radiologist assigned a PI-RADS score of 2. The lower panel shows the images assessed by mdprostate: **D:** axial T2W sequence, **E:** axial DWI sequence, and **F:** ADC map. Analysis by mdprostate includes segmentation of the prostate (green line) and identification of a diffuse hypointense area on T2W (orange line). The data at the bottom show the volume, dimensions, prostate zone, and ADC of the prostate nodule, with a PI-RADS score of 5. A prostate abscess was ruled out. There was no agreement between the radiologist (PI-RADS score 2) and mdprostate (PI-RADS score 5). Based on clinical data, antibiotic treatment was administered, and symptoms resolved. Due to persistently elevated PSA, the patient underwent prostate biopsy, which showed acute prostatitis due to *Klebsiella pneumoniae* without evidence of malignancy.

ADC: apparent diffusion coefficient; bpMRI: biparametric prostate magnetic resonance imaging; DWI: diffusion-weighted imaging; PSA: prostate-specific-antigen; PI-RADS: Prostate Imaging Reporting and Data System; PZ: peripheral zone; T2W: T2-weighted. TZ: transition zone

predominantly on the left. The DWI sequence shows mild hyperintensity. The ADC map shows marked hypointensity in the TZ and PZ of both prostatic lobes, predominantly on the left, with restricted diffusion and a mean ADC value of 762 $\mu\text{m}^2/\text{s}$. There was no agreement between the radiologist (PI-RADS score 2) and mdprostate (PI-RADS score 5). The patient underwent prostate biopsy, which showed acute prostatitis due to *Klebsiella pneumoniae* without evidence of malignancy.

Figure 4 shows a bpMRI of a 63-year-old man with a PSA of 7.14 ng/mL and a small nodule with a poorly defined border in the right posteromedial PZ. The DWI

axial sequence shows marked focal hyperintensity. The ADC map shows marked focal hypointensity (asterisk), indicating restricted diffusion with a mean ADC value of 638 $\mu\text{m}^2/\text{s}$. The radiologist assigned this nodule a PI-RADS score of 4. Mdprostate assigned the small nodule a PI-RADS 3 and segmented the nodule as larger than calculated by the radiologist. In the DWI sequence, this nodule showed marked restriction, and because it was in the peripheral zone, the radiologist assigned a PI-RADS score of 4. A targeted prostate biopsy was performed, and the histopathologic diagnosis was prostate adenocarcinoma.

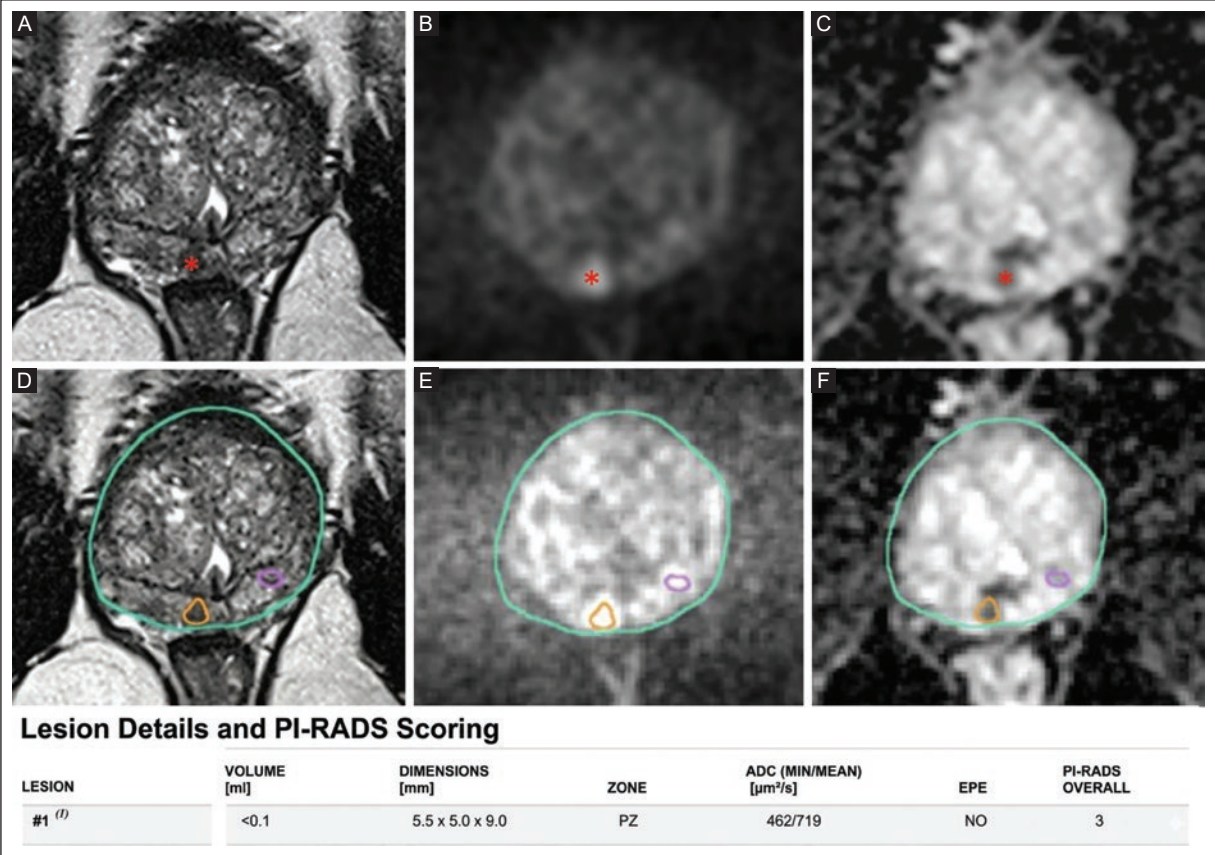


Figure 4. bpMRI of a 63-year-old man with a PSA of 7.14 ng/mL. The upper panel shows the images assessed by the radiologist: **A:** T2W axial sequence shows a nodule with poorly defined margins (asterisk) measuring 5.2 mm × 4.5 mm × 6.1 mm along the transverse, anteroposterior, and longitudinal axes, respectively, with an estimated volume of 0.07 mL. The nodule is located in the right posteromedial PZ. **B:** DWI axial sequence shows marked focal hyperintensity (asterisk). **C:** ADC map shows marked focal hypointensity (asterisk), indicating restricted diffusion, with a mean ADC of 638 $\mu\text{m}^2/\text{s}$ (range 509-917). The radiologist assigned this nodule a PI-RADS score of 4. The lower panel shows the images assessed by mdprostate: **D:** axial T2W sequence, **E:** axial DWI sequence, and **F:** ADC map. Analysis by mdprostate includes segmentation of the prostate (green line) and the nodule (orange line). The data at the bottom show the volume, dimensions, prostate zone, and ADC of the prostate nodule, with a PI-RADS score of 3. Additionally, a small nodule (purple line) was scored PI-RADS 2 by both the radiologist and mdprostate. A targeted prostate biopsy was performed, and histopathologic diagnosis was prostate adenocarcinoma, Gleason 3 + 3 (ISUP 1). The patient received HIFU treatment. Mdprostate assigned the nodule a PI-RADS 3 and segmented it as larger than the radiologist calculated. In the DWI sequence, this nodule showed marked restriction, and because it was in the peripheral zone, the radiologist assigned a PI-RADS score of 4.

ADC: apparent diffusion coefficient; bpMRI: biparametric prostate magnetic resonance imaging; DWI: diffusion-weighted imaging; HIFU: high-intensity focused ultrasound; ISUP: International Society of Urological Pathology; PSA: prostate-specific antigen; PI-RADS: Prostate Imaging Reporting and Data System; PZ: peripheral zone; T2W: T2-weighted.

DISCUSSION

Our study showed that mdprostate had high diagnostic performance compared to the radiologist using bpMRI in predicting a high or very high likelihood of csPCa based on the PI-RADS score. Interobserver agreement between the radiologist and mdprostate was almost perfect. Mdprostate can be used as a second reader for bpMRI to predict the likelihood of csPCa based on the PI-RADS score.

AI-based tools for prostate MRI demonstrate good diagnostic performance in detecting and classifying

csPCa using the PI-RADS score. Multicenter studies demonstrate that AI systems achieve diagnostic performance comparable to or better than that of experienced radiologists^{15,16}. Bayerl et al.⁸ using mdprostate and PI-RADS v2.1 in a retrospective study of 123 German men with suspected PCa, found that 53 (43.1%) did not have PCa, 15 (12.2%) had prostate cancer (Gleason score 6) and 55 (44.7%) had csPCa. PI-RADS ≥ 4 had a sensitivity of 85.5% and a specificity of 63.2%. PI-RADS 5 had a sensitivity of 88.7%. Results were comparable to those of an expert radiologist. Chen et al.¹¹ in a meta-analysis compared the performance

of various AI systems, including mdprostate versus radiologists in detecting csPCa in a multinational study with 20,423 patients from Asia, Europe, and the United States. Both 1.5T and 3.0T mpMRI scanners from various brands (Siemens, General Electric, and Philips) were used. AI systems showed diagnostic performance with a sensitivity of 87% and specificity of 61%, comparable to radiologists with a sensitivity of 85% and a specificity of 63% for detecting csPCa. In our study, mdprostate showed higher sensitivity (96.3%) than the radiologist (81.2%) in predicting the likelihood of csPCa based on PI-RADS using bpMRI, even higher than that reported by Bayerl et al.⁸ and Chen et al.¹¹. The high sensitivity observed in our study supports the clinical utility of mdprostate as a diagnostic tool in prostate bpMRI interpretation to accurately predict the likelihood of csPCa based on PI-RADS for suspicious prostate lesions.

AI software tends to overdiagnose, especially for PI-RADS 3 lesions, prostatitis, and benign prostatic hyperplasia, resulting in unnecessary biopsies¹⁷. In their meta-analysis, Chen et al.¹¹ reported that specificity is generally higher among radiologists, suggesting that increased lesion detection may lead to overclassification. The international Prostate Imaging-Cancer Artificial Intelligence (PI-CAI) project by Saha et al.¹⁵ showed that certain AI algorithms can reduce false positives in detecting prostate cancer compared with specialized radiologists. AI software achieved performance that was not inferior to, and was even superior to, that of expert radiologists using PI-RADS v2.1 for csPCa in a multicenter, retrospective study of 10,207 prostate MRI scans from 9,129 patients in the Netherlands and Norway. Patients had suspected prostate lesions due to elevated PSA, abnormal digital rectal examination, or both. In total, 2,440 cases had csPCa confirmed by biopsy. Both 1.5T and 3.0T MRI equipment were used (Siemens Healthineers and Philips Medical Systems, respectively). The AI system was trained primarily with bpMRI, while radiologists evaluated mpMRI with PI-RADS v2.1. The AI was trained to prioritize detecting csPCa. AI software, including mdprostate, showed fewer false positives than radiologists (56 versus 113, respectively). The authors emphasize the ability of AI software to improve clinical specificity by achieving high detection of csPCa, reducing false positives, and lowering the risk of overdiagnosis. In our study, mdprostate demonstrated a tendency toward higher PI-RADS categorization, particularly for PI-RADS 5 lesions. The radiologist classified 11 (13.4%) cases, while mdprostate classified 21 (25.6%) cases. Therefore, specificity was lower for mdprostate (89.1%) compared to the radiologist (98.0%). This finding is

consistent with the significant differences identified by Bowker's test and may reflect a diagnostic strategy of mdprostate to overestimate the PI-RADS score.

Previous reports have shown mild to high interobserver variability in PI-RADS classification, depending on the radiologist's experience and MRI image quality^{18,19}. Therefore, AI tools are proposed to improve the standardization and reproducibility of prostate cancer diagnosis. The multicenter study by Forookhi et al.²⁰ demonstrated that Quantib® Prostate software v2.0.1 (Quantib BV, Rotterdam, The Netherlands) improved agreement among less experienced radiologists. Cohen's Kappa showed excellent interobserver agreement between mdprostate and radiologists, using two categories: nonsuspicious (PI-RADS 1-3) and suspicious prostate cancer (4 and 5), and found improvement in PI-RADS assessment by less experienced radiologists. The multicenter PI-CAI study¹⁵ reported that AI algorithms match or exceed the average performance of radiologists in detecting csPCa. Software such as mdprostate has demonstrated high concordance ($k > 0.80$), considered almost perfect. These results are consistent with reports that AI can be effectively integrated into clinical practice as a diagnostic support system, especially for less experienced radiologists, with reductions in false positives and less interobserver variability^{8,15}. Our results showed comparable findings with almost perfect interobserver agreement ($k = 0.81$, 95% CI 0.67-0.96) between the radiologist and mdprostate in assigning PI-RADS scores. These findings support the evidence that AI-assisted prostate MRI interpretation with software such as mdprostate may serve as a reliable complementary tool to improve lesion detection, reduce diagnostic variability, and potentially facilitate more standardized reporting in clinical practice.

The main strength of our study is that it evaluated diagnostic performance and interobserver agreement of mdprostate, allowing a comprehensive assessment in a clinical setting, which is infrequent in the literature. However, the study has several limitations. First, this was a retrospective single-center study with a limited sample size, which may restrict the generalizability of the findings. Second, histopathological confirmation of prostate cancer was not available for all patients, limiting the use of the standard reference for the radiologist to evaluate the diagnostic performance of mdprostate. Third, a single radiologist performed the PI-RADS evaluation, so interobserver variability between radiologists was not assessed. Finally, two MRI scanners were used, which may have caused variability in image features.

CONCLUSION

Mdprostate demonstrated high diagnostic performance compared to the radiologist using bpMRI for the prediction of high and very high likelihood of csPCa based on the PI-RADS score. Additionally, agreement between the radiologist and mdprostate was almost perfect for PI-RADS scoring of prostate lesions. Implementing mdprostate as a second-reading tool may improve diagnostic standardization by prioritizing sensitivity over specificity⁸. Its progressive integration into clinical practice will likely play an increasingly important role in detecting and characterizing csPCa.

Acknowledgments

The authors thank Professor Ana M. Contreras-Navarro for her guidance in preparing and writing this scientific paper. This original research in the Radiology Specialty field was an awarded thesis at the Cuarta Convocatoria Nacional 2025-2026 "Las Mejores Tesis de Radiología para Publicar en el JMeXFRI".

Funding

The authors declare that they have not received funding.

Conflicts of interest

The authors declare no conflicts of interest.

Ethical considerations

Protection of human subjects and animals. The authors declare that the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the World Medical Association and the Declaration of Helsinki (1964) and subsequent amendments.

Confidentiality, informed consent, and ethical approval. The authors have followed their institution's confidentiality protocols. All participants provided written informed consent.

Declaration on the use of artificial intelligence. The authors declare that no generative artificial intelligence was used in the writing or creation of the content of this manuscript.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209-249. doi: 10.3322/caac.21660.

2. Kuhl CK, Schrading S, Enno-Edzard B, Heidenreich A, Katz-Brull R, Pickhardt PJ, et al. Abbreviated biparametric prostate MR imaging in men with elevated prostate-specific antigen. *Radiology.* 2017;285(2):493-505. doi: 10.1148/radiol.2017170129.
3. van der Leest M, Cornel E, Israël B, Hendriks R, Padhani AR, Hoogenboom M, et al. 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.* 2019;75(4):570-578. doi: 10.1016/j.eururo.2018.11.023.
4. 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(3):793-804. doi: 10.1148/radiol.2016152542.
5. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health.* 2019;1(6):e271-e297. doi: 10.1016/S2589-7500(19)30123-2.
6. Bhattacharya I, Khandwala YS, Vesal S, Shao W, Yang Q, Soerensen SJC, et al. A review of artificial intelligence in prostate cancer detection on imaging. *Ther Adv Urol.* 2022;14:17562872221128791. doi: 10.1177/17562872221128791.
7. Russo T, Quarta L, Pellegrino F, Cosenza M, Camisassa E, Lavallo S, et al. The added value of artificial intelligence using Quantib Prostate for the detection of prostate cancer at multiparametric magnetic resonance imaging. *Radiol Med.* 2025;130(7):1105-1114. doi: 10.1007/s11547-025-02017-8.
8. Bayerl N, Adams LC, Cavallaro A, Bäuerle T, Schlicht M, Wullich B, et al. Assessment of a fully-automated diagnostic AI software in prostate MRI: clinical evaluation and histopathological correlation. *Eur J Radiol.* 2024;181:111790. doi: 10.1016/j.ejrad.2024.111790.
9. Baydoun A, Jia AY, Zaorsky NG, Kashani R, Rao S, Shoag JE, et al. Artificial intelligence applications in prostate cancer. *Prostate Cancer Prostatic Dis.* 2024;27(1):37-45. doi:10.1038/s41391-023-00684-0.
10. Kim H, Kang SW, Kim JH, Nagar H, Sabuncu M, Margolis DJA, et al. The role of AI in prostate MRI quality and interpretation: opportunities and challenges. *Eur J Radiol.* 2023;165:110887. doi:10.1016/j.ejrad.2023.110887. Erratum in: *Eur J Radiol.* 2024;177:111585. doi: 10.1016/j.ejrad.2024.111585.
11. Chen H, Li E, Christos PJ, Zhu YS. Comparison of artificial intelligence and radiologists in MRI-based prostate cancer diagnosis: a meta-analysis of accuracy and effectiveness. *Biomedicine.* 2025;14(1):20. doi: 10.3390/biomedicine14010020.
12. Giganti F, Rosenkrantz AB, Villeirs G, Panebianco V, Stabile A, Emberton M, et al. The evolution of MRI of the prostate: The past, the present and the future. *AJR Am J Roentgenol.* 2019;213 (2):384-396. doi: 10.2214/AJR.18.20796.
13. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate imaging reporting and data system version 2.1: 2019 update of adherence to prostate imaging reporting and data system version 2. *Eur Urol.* 2019;76(3):340-351. doi: 10.1016/j.eururo.2019.02.033.
14. Friedrichsmeier T, and the RKWard Team (2024). RKWard: graphical interface to the statistical language R. Version 0.8.0. Available from: <https://rkwad.kde.org>
15. Saha A, Bosma JS, Twilt JJ, van Ginneken B, Bjartell A, Padhani AR, et al. Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. *Lancet Oncol.* 2024;25(7):879-887. doi: 10.1016/S1470-2045(24)00220-1.
16. Giganti F, Moreira da Silva N, Yeung M, Davies L, Fray A, Ferrer Rodriguez M, et al. AI-powered prostate cancer detection: a multi-centre, multi-scanner validation study. *Eur Radiol.* 2025;35(8):4915-4924. doi: 10.1007/s00330-024-11323-0.
17. Jensen LJ, Kim D, Elgeti T, Steffen IG, Schaafs LA, Haas M, et al. Detecting clinically significant prostate cancer in PI-RADS 3 lesions using T2w-derived radiomics feature maps in 3T prostate MRI. *Curr Oncol.* 2024;31(11):6814-6828. doi: 10.3390/curroncol31110503.
18. Wallström J, Thimansson E, Andersson J, Karlsson M, Zackrisson S, Bratt O, et al. An online national quality assessment survey of prostate MRI reading: interreader variability in prostate volume measurement and PI-RADS classification. *Eur J Radiol Open.* 2024;14:100625. doi: 10.1016/j.ejro.2024.100625.
19. Bhayana R, O'Shea A, Anderson MA, Bradley WR, Gottumukkala RV, Mojtahed A, et al. PI-RADS versions 2 and 2.1: interobserver agreement and diagnostic performance in peripheral and transition zone lesions among six radiologists. *AJR Am J Roentgenol.* 2021;217(1):141-151. doi: 10.2214/ajr.20.24199.
20. Forookhi A, Laschena L, Pecoraro M, Borrelli A, Massaro M, Dehghanpour A, et al. Bridging the experience gap in prostate multiparametric magnetic resonance imaging using artificial intelligence: a prospective multi-reader comparison study on inter-reader agreement in PI-RADS v2.1, image quality and reporting time between novice and expert readers. *Eur J Radiol.* 2023;161:110749. doi: 10.1016/j.ejrad.2023.110749.