European Journal of Radiology Open 8 (2021) 100332

Contents lists available at ScienceDirect

European Journal of Radiology Open

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

Value of an online PI-RADS v2.1 score calculator for assessment of prostate MRI

Borna K. Barth [a], Katharina Martini [a], Stephan M. Skawran [a], Florian A. Schmid [b], Niels J. Rupp [c], Laura Zuber [d], Olivio F. Donati [a,*]

[a] Institute of Diagnostic and Interventional Radiology, University Hospital Zurich and University of Zurich, Raemistrasse 100, CH-8091 Zurich, Switzerland
[b] Department for Urology, University Hospital Zürich, Rämistrasse 100, CH-8091 Zurich, Switzerland
[c] Department of Pathology and Molecular Pathology, University Hospital Zurich and University of Zurich, Raemistrasse 100, CH-8091 Zurich, Switzerland
[d] University of Zurich, Raemistrasse 78, 8006 Zurich, Switzerland

HIGHLIGHTS

- A PI-RADS Score Calculator may significantly fasten the reporting process in Prostate MRI.
- Increased PI-RADS v2.1 reporting speed does not translate into lower diagnostic accuracy.
- Particularly non-expert radiologists may profit in terms of prostate MRI reporting efficiency.

ARTICLE INFO

Keywords:
Prostate MRI
Prostate cancer
PI-RADS 2.1
Reporting software

ABSTRACT

Purpose: To evaluate the value of a browser-based PI-RADS Score Calculator (PCalc) compared to MRI reporting using the official PI-RADS v2.1 document (PDoc) for non-specialized radiologists in terms of reporting efficiency, interrater agreement and diagnostic accuracy for detection of clinically significant prostate cancer (PCa).
Methods: Between 09/2013 and 04/2015, 100 patients (median age, 64.8; range 47.5–78.2) who underwent prostate-MRI at a 3 T scanner and who received transperineal prostate mapping biopsy within <6 months were included in this retrospective study. Two non-specialized radiology residents (R1, R2) attributed a PI-RADS version 2.1 score for the most suspect (i. e. index) lesion (i) using the original PI-RADS v2.1 document only and after a 6-week interval (ii) using a browser-based PCalc. Reading time was measured. Reading time differences were assessed using Wilcoxon signed rank test. Intraclass-correlation Coefficient (ICC) was used to assess interrater agreement (IRA). Parameters of diagnostic accuracy and ROC curves were used for assessment of lesion-based diagnostic accuracy.
Results: Cumulative reading time was 32:55 (mm:ss) faster when using the PCalc, the difference being statistically significant for both readers ($p < 0.05$). The difference in IRA between the image sets (ICC 0.55 [0.40, 0.68]) and 0.75 [0.65, 0.82] for the image set with PDoc and PCalc, respectively) was not statistically significant. There was no statistically significant difference in lesion-based diagnostic accuracy (AUC 0.83 [0.74, 0.92] and 0.82 [95 % CI: 0.74, 0.91]) for images assessed with PDoc as compared to PCalc (AUC 0.82 [0.74, 0.91] and 0.74 [95 %CI: 0.64, 0.83]) for R1 and R2, respectively.
Conclusion: Non-specialized radiologists may increase reading speed in prostate MRI with the help of a browser-based PI-RADS Score Calculator compared to reporting using the official PI-RADS v2.1 document without impairing interreader agreement or lesion-based diagnostic accuracy for detection of clinically significant PCa.

1. Introduction

The Prostate Imaging and Reporting Data System (PI-RADS), developed in a joint collaboration by the European Society of Urogenital Radiology (ESUR) and the American College of Radiology (ACR), serves the purpose to standardize acquisition and interpretation of

* Corresponding author at: Institute of Diagnostic and Interventional Radiology, University Hospital Zürich, Rämistrasse 100, CH-8091 Zurich, Switzerland.
E-mail address: olivio.donati@usz.ch (O.F. Donati).

https://doi.org/10.1016/j.ejro.2021.100332
Received 30 December 2020; Received in revised form 4 February 2021; Accepted 14 February 2021
2352-0477/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

multiparametric prostate MRI. It has been evaluated and compared to non-standardized reporting [1–4] and has been revised twice, with version 2.1 [5] being the most current revision of the document. Today, PI-RADS is widely accepted in the radiologic and urologic community for reporting MRI of the prostate.

Strictly cohering to the PI-RADS rules while interpreting a prostate MRI can be challenging. First, several features which differ based on lesion location, morphology and contrast uptake kinetics have to be assessed on multiple sequences for each lesion and a diagnostic algorithm has to be followed to attain a final PI-RADS score. This process can be time-consuming. As cost of care and efficacy have increasingly become important factors in healthcare in general, reporting time has become an important performance indicator. Second, lesion morphology very often tends to be visually equivocal, presumably due to (i) a generally small size of the target and (ii) inherent technical difficulties to visualize such a lesion on MRI [6], thereby leaving room for subjectivity. At the same time 'quantification' of features and semi-quantitative metrics are not currently integrated in PI-RADS v2.1 in order to keep the reporting process as simple as possible. These factors lead to considerable variability in inter- and intrareader agreement [7–12].

Despite continuous improvement, an expert radiologist's performance using PI-RADS shows a mean sensitivity of 85 % for detection of significant PCa when following the interpretation suggestions of PI-RADS v.2.1 [13,14]. The performance tends to be inferior at low throughput centers [15].

In an attempt to simplify and speed up the reporting process of prostate MRI, we use a browser-based PI-RADS Score Calculator (PCalc) to facilitate navigation through the PI-RADS v2.1 decision tree and provide the radiologist with visual examples of the morphology of lesions for each PI-RADS score. We derived the hypothesis that for non-specialized radiologists reporting prostate MRI with the assistance of PCalc might reduce reading time and increase reproducibility without impairing the accuracy in detection of clinically significant prostate cancer (PCa).

The purpose of this study was to compare the value of a browser-based PCalc to MRI reporting using the official PI-RADS v2.1 document (PDoc) for non-specialized radiologists in terms of reporting efficiency, interrater agreement and lesion-based diagnostic accuracy for clinically significant PCa.

2. Material and methods

The regional ethics committee approved this retrospective study and written general informed consent was available from all patients prior to the examination. The study was compliant with the Health Insurance Portability and Accountability Act (HIPAA).

2.1. Study population

Between September 2013 and April 2015, 105 consecutive treatment-naïve men were included. Inclusion criterion were (i) a clinically indicated multiparametric MRI of the prostate at 3 T (Tesla) for either biopsy-proven PCa, elevated prostate specific antigen (PSA) and/or suspicious digital rectal examination, and (ii) who subsequently underwent transperineal prostate mapping biopsy (TPMB) within <6 months. Patients with an incomplete scan protocol (n = 3), patients who received TPMB more than once (n = 1) and patients with prior androgen deprivation therapy (n = 1) were excluded. The final study population comprised of 100 patients (median age, 64.8; range 47.5–78.2) with a mean PSA, 7.7 μg/L (range, 0.8–34.7).

Patient demographics and a summary of histopathology data is shown in Table 1 .

Table 1
Patient demographics and histopathology data from transperineal prostate mapping biopsy (TPMB).

Demographic Parameter	Value
Age (y)*	64.8 [47.5–78.2]
PSA at time of MRI imaging (in μg/L)*	7.7 [0.8–34.7]
Prostate volume at MRI (in ml)*	48.6 [14.9–135.5]
PSA density at time of MRI imaging*	0.19 [0.02–0.37]
Time interval between mpMRI and Histopathology (d)*	39.9 [1–161]
Patients diagnosed with clinically significant cancer**	47/100 (47)
Patients with additional TRUS biopsy <180d prior to MRI	11/100 (11)
Prevalence of Index Gleason Scores on TPMB**	
3 + 4	19/47 (40.4)
4 + 3	14/47 (29.8)
4 + 4	5/47 (10.6)
4 + 5	9/47 (19.1)

Note: n = 100 patients are in the dataset.
* Data are means; data in squared parentheses are ranges.
** Data are absolute counts; data in parentheses are percentages.

2.2. Technical background

Images were acquired on 3 T MRI systems (MAGNETOM Skyra, SIEMENS Healthcare®, Erlangen, Germany) equipped with two independent transmit channels (TimTX TrueShape, Siemens Healthcare, Erlangen, Germany). An 18-channel phased-array coil and a balloon-covered, expandable endorectal coil (ERC; Medrad, Warrendale, Pa) were used as receiver coils.

T2-weighted turbo-spin echo images (T2w) were obtained in three orthogonal planes (transversal, sagittal and coronal), covering the whole prostate gland including the seminal vesicles. Diffusion-weighted images (DWI) were acquired in the transverse plane with identical orientation as the T2w images. Apparent diffusion coefficient (ADC) parametric maps were calculated using a mono-exponential fit based on three obtained b-values (either 0-50-1000s/mm^2 [n = 86] or 100, 600, 1000s/mm^2 [n = 14]). A high-b-value (1400s/mm^2) was calculated based on a standard mono-exponential fit. Dynamic contrast enhanced images (DCE-MRI) were obtained in transverse sections using a 3D T1w spoiled gradient-echo pulse sequence with a temporal resolution of <8 s. The contrast agent Gadoterate meglumine (Dotarem; Guerbet, Darmstadt, Germany) was used with a dosage of 0.1 mmol/kg bodyweight. Contrast injection was performed with an automated MR injection system (Spectris Solaris EP, MEDRAD MR Injector, Bayer HealthCare LCC, Whippany NJ) at a flow rate of 2 mL/sec.

T1-weighted turbo-spin echo images (T1w) were additionally obtained during the clinical routine but were not used in this study setup. Basic scan parameters of MRI sequences used for study purposes are shown in Table 2.

2.3. PI-RADS score calculator (PCalc)

The browser-based PCalc used in this study is freely accessible [16].

Table 2
Basic scan parameters of MRI sequences used for study purposes.

	T2w	T2w	DWI	DCE-MRI
Number of averages	3	1/2	2, 4, 8	–
TR range (ms)	3260–5140	3130–4000	4900–5718	5.0–6.3
TE range (ms)	93–97	97–102	69–96	1.8
In-plane resolution (mm)	0.27 × 0.27	2.5 × 2.4	0.7 × 0.7	1 × 0.6
Slice thickness (mm)	3	3	3	3
Acquisition Plane	Transverse	Coronal/ Sagittal	Transverse	Transverse

It supports the user by providing step-by-step guidance throughout the assessment process, including visual examples of each PI-RADS lesion category in order to facilitate assigning a final PI-RADS score.

Reporting pathway:

1 First the user needs to address whether the lesion is located within the peripheral (PZ) or the transitional zone (TZ).
2 For the PZ the user first has to select a DWI/ADC appearance of the lesion. Visual examples and the official PI-RADS v2.1 text descriptions of a specific category are given. If the lesion appears to have a PI-RADS 3 DWI morphology, the user has to assess the DCE-MRI additionally in order to decide whether the lesion is 'DCE positive' or 'DCE negative' according to PI-RADS v2.1.
3 For the TZ the user first has to select a T2w appearance. If the lesion appears to have a PI-RADS 2 or 3 T2w morphology, the user has to assess the DWI/ADC appearance of the lesion in order to get a final score.

2.4. Reference standard

All patients underwent transperineal prostate mapping biopsy (TPMB) performed within <6 months (mean [d], 39.9, range 1−161) after MRI. TPMB was performed under general anesthesia using a 5-mm brachytherapy transperineal template grid. Histopathology specimens were evaluated by dedicated genitourinary pathologists. The histopathology results were documented (i) within an official, text-based document and (ii) on a TPMB report scheme, the latter allowing for visual correlation of the lesion in relation to the segmental anatomy of the gland. The report included the number, location and length of the lesions. Lesions with a Gleason score (GS) ≥3 + 4 were considered as clinically significant. If multiple lesions were present, the one with the highest GS was considered as index lesion.

2.5. Readout

Images were anonymized. One staff radiologist (initials blinded for review) with 6 years of experience in prostate MRI reporting identified the index lesion (if present) on the TPMB report schemes. After visual side-by-side matching of the pathology data with the MRI images, the radiologist marked the index lesion with an arrow (on either a T2w or DWI image, depending on where the lesion was depicted best) and stored the image as a screenshot in the corresponding patient's anonymized image folder. In order to generate a pool of true negative lesions if pathology revealed no clinically significant tumor, the prostate of patients without PCa was divided into quadrants. For each patient (i) a quadrant (i. e. anterior left [1], posterior left [2], anterior right [3], posterior right [4]) was selected by using a randomizer tool which generated a random sequence of numbers from 1 to 4. Within each selected quadrant (ii) the radiologist (initials blinded for review) chose one lesion/morphologic finding and marked it with an arrow.

Two radiology residents [R1] (initials blinded for review) with 4 years of experience and [R2] (initials blinded for review) with 2 years of experience in reporting prostate MRIs under senior supervision) independently performed the image readout in two separate sessions. Both readers had previously used the updated PI-RADS System within clinical routine (>50 prostate MRs) and thus were familiar with it. The readers were aware that patients received TPMB after MRI, but were blinded to the remaining clinical information. In both sessions the MRI scans were presented in random order. The first readout was performed without using the PCalc. Readers were instructed to use the official PI-RADS v2.1 document [5] for assistance if they deemed it to be necessary. The second readout of the same data set was performed after a time interval >6 weeks and with new randomization codes in order to minimize recall bias. This time, the readers were instructed to use the PCalc for every lesion side-by-side on a separate monitor. The use of the PCalc was mandatory. The readers were instructed to assess the areas marked by an arrow on the corresponding screenshot according to the PI-RADS v2.1 criteria [5]. Apart from the PI-RADS Score the readers had to measure the reading time starting with the actual reading process (i.e. when all necessary images were loaded and ready for interpretation) and ending with the final assessment of the lesion score. The readout was performed using an electronic reader form.

2.6. Statistical analysis

Demographic variables were summarized as means with ranges and absolute numbers with percentages. Categorical variables were summarized as absolute figures. Kolmogorov-Smirnov test was used to assess the distribution of data.

Wilcoxon signed rank test was used to evaluate the central tendency between PI-RADS Scores and reading time.

Agreement for the attributed PI-RADS scores (inter- and intrareader) was assessed using intraclass-correlation coefficient (ICC, two-way random, single measures), including 95 % confidence intervals (CI). The ICC values were interpreted as follows: 0–0.40, poor agreement; 0.41–0.58, fair agreement; 0.59–0.75, good agreement; 0.76–1.0, excellent agreement [17].

Based on binarized data for the presence of clinically significant PCa (PI-RADS scores 1–3 were deemed to be 'negative' and scores 4 and 5 'positive'), the area under the receiver-operating characteristics curve (AUC) was calculated for the according PI-RADS scores attained with PDoc and with PCalc. Measures of diagnostic performance - sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) - were calculated. CIs for sensitivity, specificity and accuracy are presented as "exact" Clopper-Pearson confidence intervals. CIs for the positive and negative predictive values are the standard logit confidence intervals (CIs) [18]. Statistical significance was defined as non-overlapping 95 % CIs [20]. A p-value of <0.05 was considered statistically significant.

Statistical analysis was performed with IBM SPSS statistical software (SPSS version 21; Chicago, Ill).

3. Results

3.1. Patient demographics and histopathology data

Forty-seven of 100 patients (47 %) had clinically significant PCa on TPMB. The most prevalent Gleason score (GS) for an index lesions was 3 + 4, diagnosed in n = 19/47 (40.4 %) patients, followed by a GS of 4 + 3, diagnosed in n = 14/47 (29.8 %), a GS of 4 + 5, diagnosed in n = 9/47 (19.1 %) and a GS of 4 + 4, diagnosed in n = 5/47 (10.6 %). Patient demographics and histopathology data from transperineal prostate mapping biopsy (TPMB) are shown in Table 1 .

3.2. Reporting speed per lesion

Average reading time was 35 s [range, 13–90] for R1 and 25 s [range, 4–75] for R2, as compared to 21 s [range, 5–56] for R1 and 20 s [range, 4–53] for R2 when reading with PDoc and PCalc, respectively. Cumulative reading time was shorter when using PCalc for both readers (Time difference R1: -23:54 [mm:ss], and R2: -09:01 [mm:ss]). Difference in reading time was statistically significant for both readers ($p < 0.05$). Reading time per lesion is shown in Fig. 1, two lesion examples are shown in Fig. 3.

R1 seemed to shorten reading time by using PCalc particularly when attributing PI-RADS 1, 2 and 3 scores to lesions (% range of mean time saved/lesion, 43.8–47.1 %), whereas R2 seemed to profit mainly for PI-RADS 1 lesions (% of mean time saved/lesion, 42.9 %) when using PCalc. Percentage (%) of mean time saved per PI-RADS score is shown in Table 3.

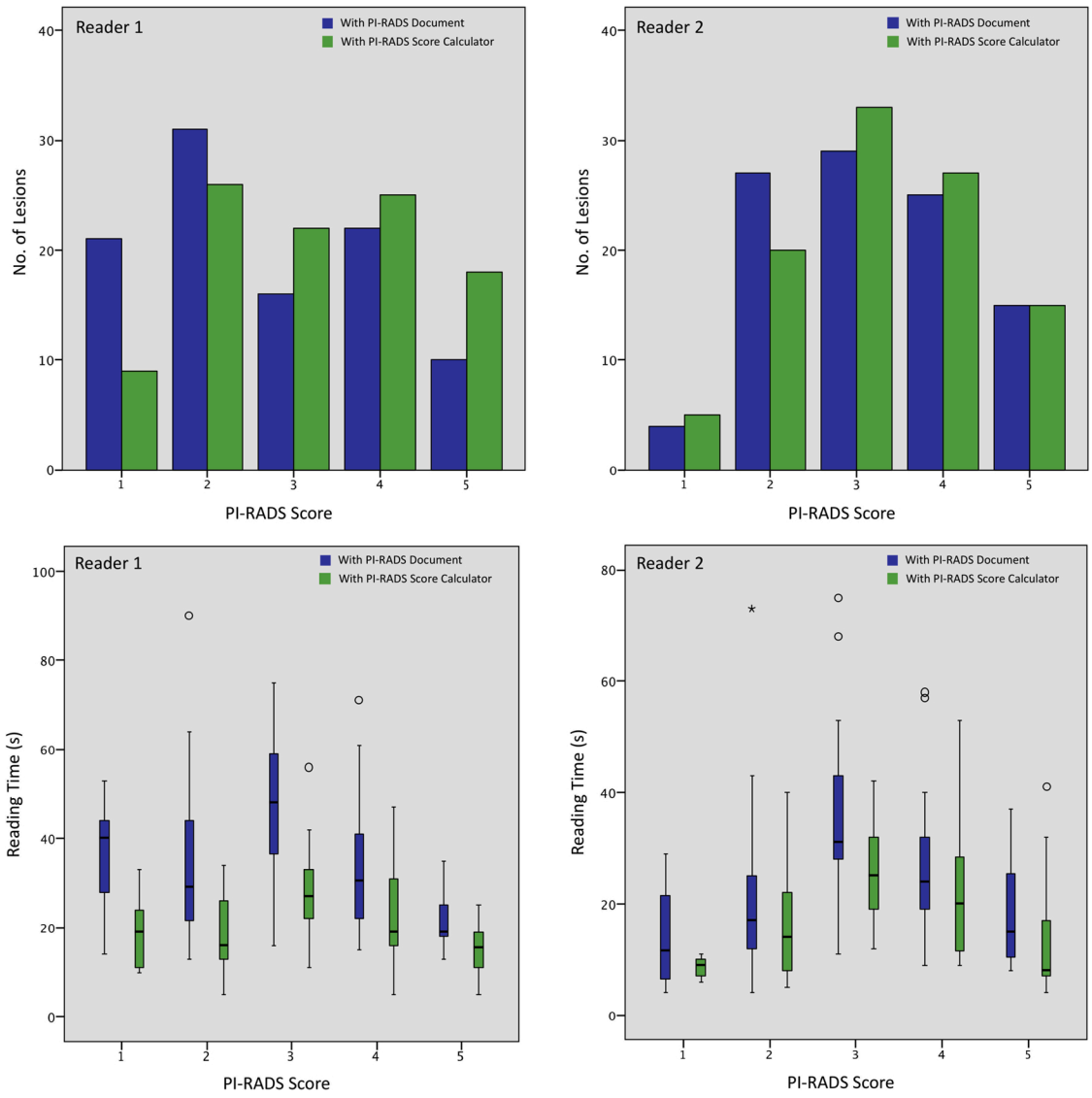

Fig. 1. PI-RADS Scores and Reading Time.
PI-RADS v2.1 scores attributed to the lesions for images acquired with PI-RADS Document and with PI-RADS Score Calculator (upper row) and reading time per lesion (lower row) are shown.

Table 3
Percentage (%) of mean time saved per PI-RADS Score using PI-RADS Score Calculator.

PI-RADS	1	2	3	4	5	Mean
Reader 1*	45.7	47.1	43.8	30.3	28.6	39.1
Reader 2*	42.9	20.0	16.7	15.4	17.6	22.5

Note: n = 100 patients are in the dataset.
* Data are percentages.

3.3. Agreement

Interreader agreement between R1 and R2 regarding PI-RADS score was poor (ICC 0.56 [0.40, 0.68]) for the image set interpreted with PCalc and fair (ICC 0.69 [0.43, 0.82] for the image set with PDoc, respectively.

Intrareader agreement regarding PI-RADS score was excellent for R1 (ICC 0.85 [0.78, 0.90]) and R2 (ICC 0.81 [0.72, 0.87]).

3.4. Lesion-based diagnostic accuracy

Mean PI-RADS score for the analyzed data set was 2.7 and 3.2 ($p < 0.001$) for R1, and 3.2 and 3.3 ($p = 0.45$) for R2, for images read using PDoc and with PCalc. PI-RADS v2.1 scores attributed to the lesions for images using PDoc and PCalc are shown in Fig. 1. An example for the tendentially higher scoring using PCalc is shown in Fig. 4.

$AUC_{Reader1}$ was 0.83 [95 %CI: 0.74, 0.92] using PDoc and 0.86 [95 % CI: 0.82, 0.95] using PCalc. $AUC_{Reader2}$ was 0.82 [95 %CI: 0.74, 0.91] using PDoc and 0.74 [95 %CI: 0.64, 0.83] using PCalc. ROC analysis revealed no statistically significant difference in diagnostic performance for both readers between images read using PDoc compared to PCalc with overlapping 95 %CIs. ROC curves for both readers with PDoc and PCalc are shown in Fig. 2.

For R1, sensitivities were 64 % [95 %CI 49, 77] and 77 % [95 %CI 62, 88] for PDoc and PCalc, respectively. For R2 the according sensitivities were 70 % [95 %CI 55, 83] and 64 % [95 %CI 49, 77]. Specificites were R1, 96 % [95 %CI 87, 100] and 87 % [95 %CI 75, 95] for PDoc and PCalc for R1 and 87 % [95 %CI 75, 95] and 77 % [95 %CI 64, 88] vor R2. Differences were not statistically significant with overlapping 95 %CIs. Index-lesion based diagnostic accuracy parameters of images read using

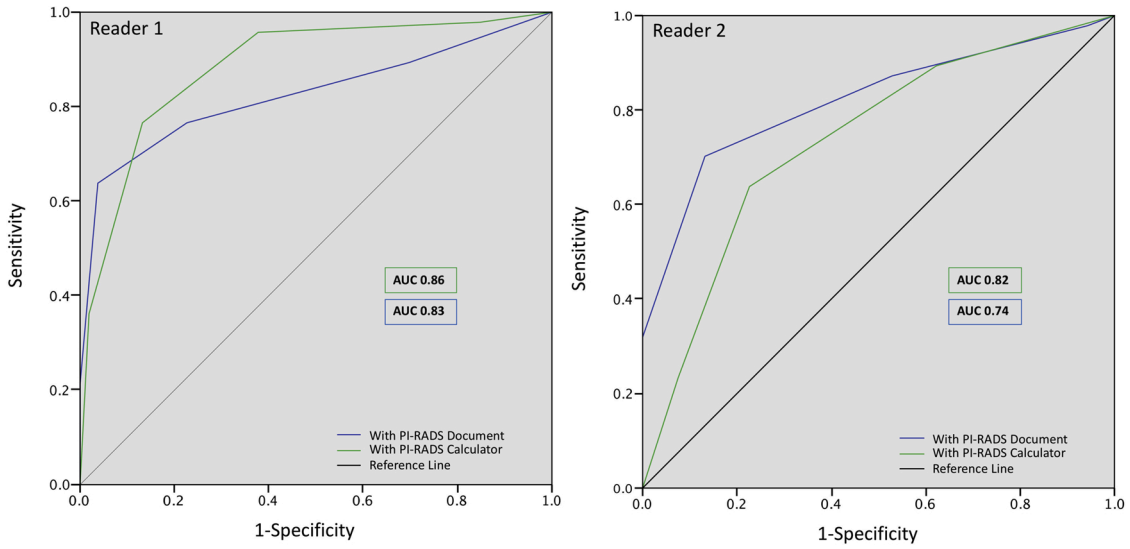

Fig. 2. Diagnostic Accuracy.
Receiver operating characteristic (ROC) curves for both readers with PI-RADS Document and with PI-RADS Score Calculator are shown.

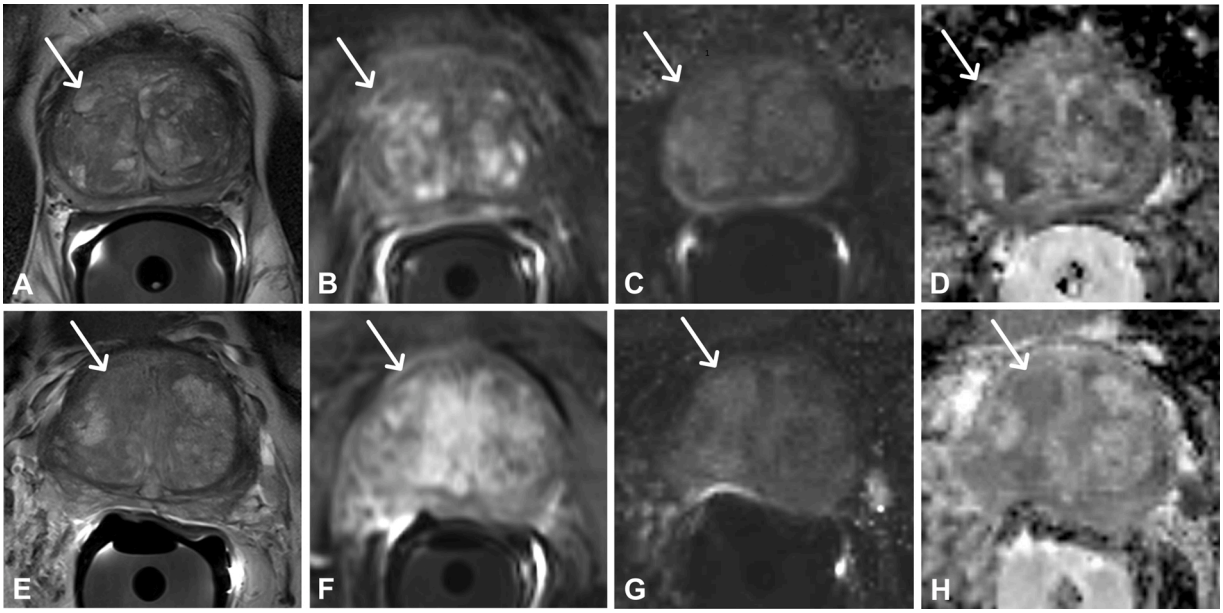

Fig. 3. Two Benign Lesions.
Two examples of definitely benign lesions within the transitional zone (TZ) in men undergoing mpMRI of the prostate for elevated PSA.
67-year-old man (A-D) demonstrating a completely encapsulated, T2w hyperintense lesion (A, T2w axial; B, DCE-MRI; C, DWI [b-1000 value]; D, ADC Map). The lesion was scored PI-RADS 1 by both readers, using PDoc and PCalc. R1 saved 5 s, R2 23 s when reading with PCalc.
54-year-old man (E-H) demonstrating a partially obscured, T2w heterogeneous, centrally hypointense lesion, with an indistinct hypointense area on ADC (A, T2w axial; B, DCE-MRI; C, DWI [b-1000 value]; D, ADC Map). The lesion was scored PI-RADS 3 by both readers with PDoc and with PCalc. R1 saved 52 s, while R2 needed 2 s longer when reading with PCalc.

PDoc and PCalc are shown in Table 4.

4. Discussion

Prostate Imaging and Reporting Data System (PI-RADS) is widely used for reporting prostate MRI. However, using PI-RADS based on the officially available document may be challenging and time-consuming, particularly for non-specialized radiologists. This may potentially impair the even wider application of this reporting algorithm in daily clinical routine. The results of this study suggest that the use of a simple, browser-based tool, which supports the radiologist to navigate along the diagnostic pathway to reach a final PI-RADS score, bears the potential to speed up reporting, without impairing interreader agreement or diagnostic accuracy.

A software tool helping to find one's way through a multi-branched decision tree in attaining a certain PI-RADS score may increase reading speed and reduce errors in misclassifying lesions due to eliminating wrong or irrelevant pathways. Using the PCalc both readers saved time by reporting approximately 30 % faster on average. The average percentage of time saved decreased with the severity of the score for both readers. Presenting only the necessary information at a given point within the decision process, as provided by PCalc, presumably accelerates the decision process. Our results are in line with a study demonstrating increased reporting efficiency in screening chest radiographs for

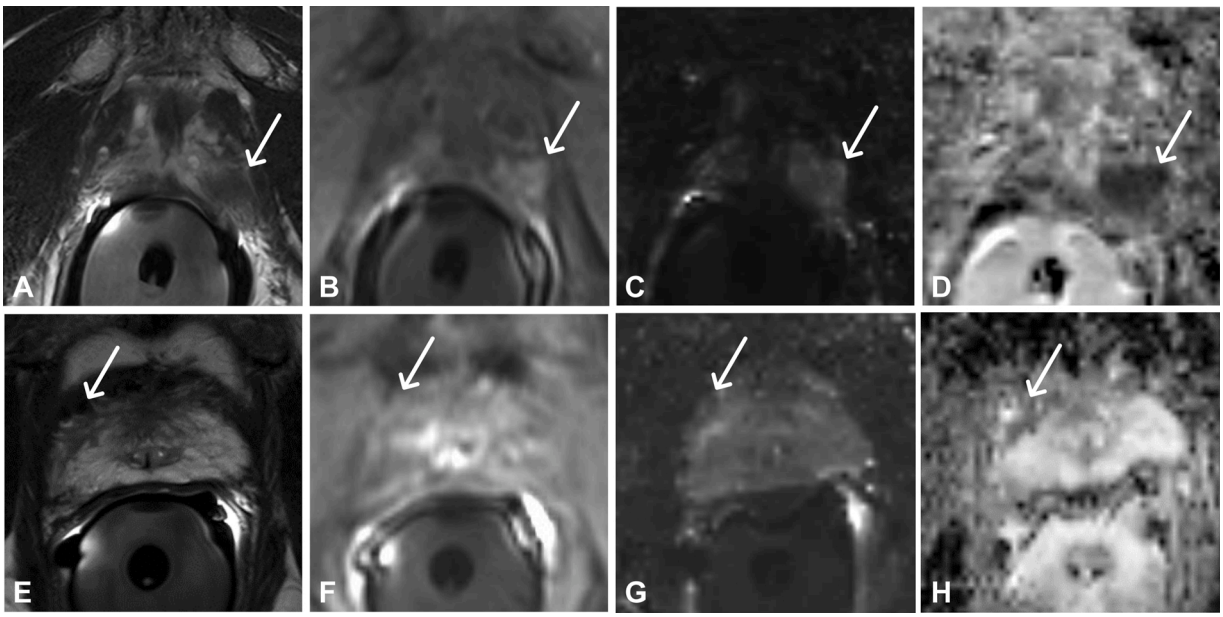

Fig. 4. Two Tumorous Lesions.
Two examples of tumorous lesions (both GS 3 + 4) within the peripheral zone (PZ) of men undergoing mpMRI of the prostate for elevated PSA.
74-year-old man (A-D) demonstrating a focal, markedly hypointense lesion on ADC with hyperintense signal on the b-1000 image (A, T2w axial; B, DCE-MRI; C, DWI [b-1000 value]; D, ADC Map). The lesion was scored PI-RADS 5 by both readers using PDoc and PI-RADS 4 by R1 and PI-RADS 5 by R2 using PCalc.
58-year-old man (E-H) demonstrating a hypointense lesion on ADC, with irregular margins (non-focal), abutting the prostate capsule (A, T2w axial; B, DCE-MRI; C, DWI [b-1000 value]; D, ADC Map). The lesion was scored PI-RADS 2 by both readers with PDoc and PI-RADS 3 with PCalc. Although scoring falsely negative, the case demonstrates the tendentially more offensive scoring using PCalc.

Table 4
Index-lesion based diagnostic accuracy of images read without and with PI-RADS Score Calculator.

Dg Acc Parameter	Reader 1		Reader 2	
	PDoc†	PCalc	PDoc†	PCalc
True Positives	30	36	33	30
True Negatives	51	46	46	41
False Positives	2	7	7	12
False Negatives	17	11	14	17
Sensitivity	64 (49, 77)	77 (62, 88)	70 (55, 83)	64 (49, 77)
Specificity	96 (87, 100)	87 (75, 95)	87 (75, 95)	77 (64, 88)
Positive Predictive Value*	94 (79, 98)	84 (72, 91)	83 (70, 91)	71 (59, 81)
Negative Predictive Value*	75 (67, 81)	81 (71, 88)	77 (68, 84)	71 (62, 78)
Accuracy*	81 (72, 88)	82 (73, 89)	79 (70, 87)	71 (61, 80)

Note: Values are expressed as probabilities (%).
* Data in brackets are 95% Confidence Intervals (CI). Estimates dependent on disease prevalence. Prevalence of clinically significant cancer within the dataset: 47 %.
† Conventional PI-RADS v2.1 Document.

tuberculosis using a custom software for computer-aided reporting [19].

As PI-RADS in its current version is still based on freely interpretable descriptions and does not include quantitative measures, there is inherent subjectivity in grading prostate lesions. The more experienced a radiologist becomes in judging lesions on MRI of the prostate, the more she or he may be bending the strict PI-RADS interpretation rules in order for a certain lesion to fall into a given category of likelihood for PCa. For example, even if a definite score is given according to strict application of the PI-RADS criteria, sometimes the lesion from a personal perspective of the radiologist seems to be more or less aggressive, particularly if the clinical or biochemical information provided by the referring clinician points into a certain direction. This in turn can tempt the radiologist to interpret lesion features in a way, which favor the initial 'intuition'. On one hand, including 'intuition' into interpretation of MRI may be beneficial as it allows for the radiologist's experience to influence lesion categorization, on the other hand non-expert readers will most likely benefit from strict adherence to the PI-RADS algorithm. Especially for the latter group (for example resident doctors and radiologists not working in high volume centers) a software-based tool for interpretation of prostate MR like PCalc may be helpful. We therefore expected the interreader agreement to improve by use of the PCalc which however was not the case. The difference in interreader agreement did not significantly differ between the two imaging interpretation workflows and the degree of agreement achieved in this study is comparable with the results of other studies using PI-RADS reporting [7–12]. A lack of improvement in interreader agreement in our study suggests that the inherent subjectivity in rating prostatic lesions is not eliminated just by adhering to a tool such as the PCalc. The introduction of quantitative measures, such as ADC may increase interreader agreement. Several studies have shown the potential of this metric in assessing the aggressiveness in prostate tumors and discriminating clinically significant from non-significant cancer [20,21]. Furthermore, ADC presents a reproducible metric even regarding MR system of different vendors [22]. However, an algorithm defined by a set of rules that is ideally supposed to be adopted by a wide range of radiologists must be simple and easily comprehensible. The evolution of the PI-RADS algorithms over time showed a trend away from quantification, as semiquantitative contrast kinetic parameters derived from DCE-MRI and spectroscopy have been removed in PI-RADS 2.0. This strategy introduces more reporting flexibility based on 'intuition', however at the cost of reproducibility. To increase simplicity along with agreement will be one the major challenges in future PI-RADS versions. A software-based tool may be a step towards achieving those seemingly contradictory goals and could be included in a further version of the PI-RADS document.

Diagnostic accuracy did not change significantly between the two interpretation workflows. Lesion-based sensitivity for detection of PCa

ranged between 64–77 %, which is below an average of 85 %- a value derived form a representative meta-analysis [13] assessing diagnostic accuracy using PI-RADS v2. Probably the lack of reporting experience of the two readers who were resident radiologists explains this difference in sensitivity as compared to the expert reader's performance in the above-mentioned meta-analysis.

Our study had limitations, first, we chose two non-expert readers for this study setup, as we assumed that particularly those radiologists would profit from this tool. Therefore, the results may not be applicable to experts or more experienced senior radiologists. However, the primary aim of this study was to investigate the difference in reporting time as a measure of efficacy when using the PI-RADS algorithm and inter-reader agreement, an issue which is especially pronounced in younger inexperienced radiologists and which is relevant in teaching hospitals with resident or fellowship programs. Second, the methodology for inclusion of 'truly negative' lesions is subject to a selection bias. However, along with the topographic restriction for lesion selection, the image set was identical for both readouts, without and with PCalc, hence (if present) the bias was affecting both in the same way. Moreover, diagnostic accuracy was assessed on a lesion basis, hence the results must be interpreted with caution when being compared with cancer detection data on a patient level.

In conclusion, the present study demonstrates that reporting speed in prostate MRI may be increased by using a browser-based PI-RADS Score Calculator as compared to reporting using the official PI-RADS v2.1 document without impairing interreader agreement or diagnostic accuracy for detection of clinically significant PCa. Providing such a software tool in a revised version of the PI-RADS document may further increase the wide distribution of this reporting algorithm and may simplify its application.

Author's contribution

Borna K. Barth, MD: Study Idea, Manuscript writer, Statistics.
Katharina Martini, MD: Reader 1 (Qualitative Readout).
Stephan M. Skawran, MD: Reader 2 (Qualitative Readout).
Florian A. Schmidt, MD: Urologic Input, Editing.
Niels J. Rupp, MD: Histopathology Reports and Editing.
Laura Zuber, Collection collection and postprocessing of readout data.
Olivio F. Donati, MD: Study Idea, Editing, Supervising.

Ethical statement

The regional ethics committee approved this retrospective study and written general informed consent was available from all patients prior to the examination. The study was compliant with the Health Insurance Portability and Accountability Act (HIPAA).

Funding

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

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A.B. Rosenkrantz, S. Kim, R.P. Lim, et al., Prostate cancer localization using multiparametric MR imaging: comparison of prostate imaging reporting and data system (PI-RADS) and likert scales, Radiology 269 (2) (2013) 482–492.
[2] R. Renard-Penna, P. Mozer, F. Cornud, et al., Prostate imaging reporting and data system and likert scoring system: multiparametric MR imaging validation study to screen patients for initial biopsy, Radiology 275 (2) (2015) 458–468.
[3] T. Barrett, A. Rajesh, A.B. Rosenkrantz, P.L. Choyke, B. Turkbey, PI-RADS version 2.1: one small step for prostate MRI, Clin Radiol. 74 (11) (2019) 841–852.
[4] E. NiMhurchu, F. O'Kelly, I.G. Murphy, et al., Predictive value of PI-RADS classification in MRI-directed transrectal ultrasound guided prostate biopsy, Clin. Radiol. 71 (4) (2016) 375–380.
[5] B. Turkbey, A.B. Rosenkrantz, M.A. Haider, et al., Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2, Eur. Urol. 76 (3) (2019) 340–351.
[6] P.J. van Houdt, G. Ghobadi, I.G. Schoots, et al., Histopathological features of MRI-invisible regions of prostate cancer lesions, J. Magn. Reson. Imaging 51 (4) (2020) 1235–1246.
[7] K.M. Selnaes, A. Heerschap, L.R. Jensen, et al., Peripheral zone prostate cancer localization by multiparametric magnetic resonance at 3 T: unbiased cancer identification by matching to histopathology, Invest Radiol. 47 (11) (2012) 624–633.
[8] O.F. Donati, S.I. Jung, H.A. Vargas, et al., Multiparametric prostate MR imaging with T2-weighted, diffusion-weighted, and dynamic contrast-enhanced sequences: are all pulse sequences necessary to detect locally recurrent prostate cancer after radiation therapy? Radiology 268 (2) (2013) 440–450.
[9] J.J. Futterer, A. Briganti, P. De Visschere, et al., Can clinically significant prostate cancer Be detected with multiparametric magnetic resonance imaging? A systematic review of the literature, Eur. Urol. 68 (6) (2015) 1045–1053.
[10] M. Kasel-Seibert, T. Lehmann, R. Aschenbach, et al., Assessment of PI-RADS v2 for the detection of prostate cancer, Eur. J. Radiol. 85 (4) (2016) 726–731.
[11] W.C. Lin, V.F. Muglia, G.E. Silva, S. Chodraui Filho, R.B. Reis, A.C. Westphalen, Multiparametric MRI of the prostate: diagnostic performance and interreader agreement of two scoring systems, Br. J. Radiol. 89 (1062) (2016), 20151056.
[12] A.S. Becker, A. Cornelius, C.S. Reiner, et al., Direct comparison of PI-RADS version 2 and version 1 regarding interreader agreement and diagnostic accuracy for the detection of clinically significant prostate cancer, Eur. J. Radiol. 94 (2017) 58–63.
[13] L. Zhang, M. Tang, S. Chen, X. Lei, X. Zhang, Y. Huan, A meta-analysis of use of prostate imaging reporting and data system version 2 (PI-RADS V2) with multiparametric MR imaging for the detection of prostate cancer, Eur. Radiol. 27 (12) (2017) 5204–5214.
[14] A.P. Kirkham, P. Haslam, J.Y. Keanie, et al., Prostate MRI: who, when, and how? Report from a UK consensus meeting, Clin. Radiol. 68 (10) (2013) 1016–1023.
[15] K. Kohestani, J. Wallstrom, N. Dehlfors, et al., Performance and inter-observer variability of prostate MRI (PI-RADS version 2) outside high-volume centres, Scand. J. Urol. 53 (5) (2019) 304–311.
[16] Donati Of. prostate-mri.ch/pirads-calculator [Internet]. Zurich: University Hospital Zurich Available from: https://www.prostate-mri.ch.
[17] D.V. Cicchetti, S.A. Sparrow, Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior, Am. J. Ment. Defic. 86 (2) (1981) 127–137.
[18] N.D. Mercaldo, K.F. Lau, X.H. Zhou, Confidence intervals for predictive values with an emphasis to case-control studies, Stat. Med. 26 (10) (2007) 2170–2183.
[19] M. Morris, B. Saboury, N. Bandla, et al., Computer-aided reporting of chest radiographs: efficient and effective screening in the value-based imaging era, J. Digit Imaging 30 (5) (2017) 589–594.
[20] O.F. Donati, Y. Mazaheri, A. Afaq, et al., Prostate cancer aggressiveness: assessment with whole-lesion histogram analysis of the apparent diffusion coefficient, Radiology 271 (1) (2014) 143–152.
[21] O.F. Donati, A. Afaq, H.A. Vargas, et al., Prostate MRI: evaluating tumor volume and apparent diffusion coefficient as surrogate biomarkers for predicting tumor Gleason score, Clin. Cancer Res. 20 (14) (2014) 3705–3711.
[22] O.F. Donati, D. Chong, D. Nanz, et al., Diffusion-weighted MR imaging of upper abdominal organs: field strength and intervendor variability of apparent diffusion coefficients, Radiology 270 (2) (2014) 454–463.