BJR https://doi.org/10.1259/bjr.20210458

Received: 10 April 2021 | Revised: 03 October 2021 | Accepted: 05 October 2021

© 2022 The Authors. Published by the British Institute of Radiology under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License http://creativecommons.org/licenses/by-nc/4.0/, which permits unrestricted non-commercial reuse, provided the original author and source are credited.

Cite this article as:
Garmer M, Karpienski J, Groenemeyer DHW, Wagener B, Kamper L, Haage P. Structured reporting in radiologic education – Potential of different PI-RADS versions in prostate MRI controlled by in-bore MR-guided biopsies. *Br J Radiol* 2022; **95**: 20210458.

INNOVATIONS IN PROSTATE CANCER SPECIAL FEATURE : FULL PAPER

Structured reporting in radiologic education – Potential of different PI-RADS versions in prostate MRI controlled by in-bore MR-guided biopsies

[1,2]MARIETTA GARMER, MD, [1]JULIA KARPIENSKI, [1,3]DIETRICH HW GROENEMEYER, [4]BIRGIT WAGENER, [1,2]LARS KAMPER and [1,2]PATRICK HAAGE

[1]Witten/Herdecke University, Witten, Germany
[2]Clinical Radiology, Helios University Hospital Wuppertal, Wuppertal, Germany
[3]Grönemeyer Institute of Microtherapy, Bochum, Germany
[4]Medical Center – Urology, Bochum, Germany

Address correspondence to: Dr Marietta Garmer
E-mail: *marietta.garmer@uni-wh.de*

Objectives: To evaluate the efficiency of structured reporting in radiologic education – based on the example of different PI-RADS score versions for multiparametric MRI (mpMRI) of the prostate.
Methods: MpMRI of 688 prostate lesions in 180 patients were retrospectively reviewed by an experienced radiologist and by a student using PI-RADS V1 and V2. Data sets were reviewed for changes according to PI-RADS V2.1. The results were correlated with results obtained by MR-guided biopsy. Diagnostic potency was evaluated by ROC analysis. Sensitivity, specificity and correct-graded samples were evaluated for different cutpoints. The agreement between radiologist and student was determined for the aggregation of the PI-RADS score in three categories. The student's time needed for evaluation was measured.
Results: The area under curve of the ROC analysis was 0.782/0.788 (V1/V2) for the student and 0.841/0.833 (V1/V2) for the radiologist. The agreement between student and radiologist showed a Cohen's weighted κ coefficient of 0.495 for V1 and 0.518 for V2. Median student's time needed for score assessment was 4:34 min for PI-RADSv1 and 2:00 min for PI-RADSv2 ($p < 0.001$). Re-evaluation for V2.1 changed the category in 1.4% of all ratings.
Conclusion: The capacity of prostate cancer detection using PI-RADS V1 and V2 is dependent on the reader's experience. The results from the two observers indicate that structured reporting using PI-RADS and, controlled by histopathology, can be a valuable and quantifiable tool in students' or residents' education. Herein, V2 was superior to V1 in terms of inter- observer agreement and time efficacy.
Advances in knowledge: Structured reporting can be a valuable and quantifiable tool in radiologic education. Structured reporting using PI-RADS can be used by a student with good performance. PI-RADS V2 is superior to V1 in terms of inter -observer agreement and time efficacy.

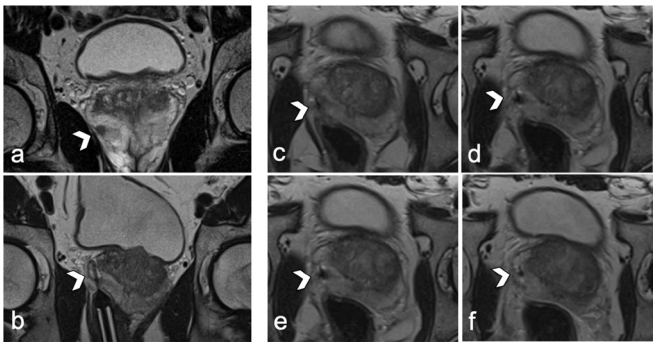
INTRODUCTION

Structured reporting with standardization of radiological reports ensures a complete analysis with higher quality and less ambiguity.[1] In combination with templates and digital tools, the results are better comparable and can be used for automatic functions. Structured reporting is well known in breast imaging (BI-RADS - Breast Imaging Reporting and Data System) and in prostate imaging (PI-RADS - Prostate Imaging Reporting and Data System). Further reporting systems exist for hepatocellular carcinoma (LI-RADS), thyroid nodules (TI-RADS) and CT diagnosis of coronary artery disease.[2–4]

Garcia-Reyes et al evaluated a reader education program on detection of prostate cancer and showed a significant increase in diagnostic accuracy and confidence for index lesion detection in a small population.[5] Structured reports might contribute to effective training and supervision of residents and students. Reproducibility is expected to improve diagnostic accuracy, however, simplification could be detrimental.[6,7] Simultaneously, an excessive increase of the time needed for assessment should be avoided.

The incidence of prostate cancer has declined or stabilized in most of the countries in Northern and Western Europe.

Figure 1. In-bore MRI-guided prostate biopsy, documentation of needle course; patient 68 years; prostate-specific antigen PSA 8.4 ng ml^{-1}; prostate volume 65 ml, PSA density 0.13 ng/ml^2; a: coronal T2-weighted diagnostic imaging with small lesion on the right side; b: coronal and c-f: consecutive axial fast T2-weighted biopsy imaging showing the needle position inside the lesion.



This trend may reflect updated recommendations for limited use of PSA screening. Nevertheless, the incidence is still high. Prostate cancer represents the second most diagnosed cancer among males worldwide.[8] In 2012, the ESUR presented guidelines for magnetic resonance imaging (MRI) in prostate cancer including the structured reporting system PI-RADS.[9] Multiple components of MRI as T_2-weighted imaging (T2w), diffusion-weighted imaging (DWI) and dynamic contrast enhanced imaging (DCE) are scored separately on a five-point Likert scale. The published criteria for this classification are based on consensus expert opinion and literature evidence. The combination of these separate scores yields an overall PI-RADS score. Similar to the known BI-RADS score for radiologic breast diagnostic, this score represents the probability for clinically significant cancer, that is unlikely to be present with PI-RADS one and highly likely to be present with PI-RADS 5.

PI-RADS Version 1 (V1) was updated in 2015 as PI-RADS V2.[10,11] In particular, the correlation of a lesion to the prostate zonal anatomy was considered and the evaluation of DCE was simplified. Minor changes were released in 2019 as PI-RADS V2.1.[12] Several comparative studies are dealing with the accuracy and inter-observer agreement of different PI-RADS versions.[13–27] Data are limited in the literature on the efficiency of PI-RADS in education of readers without specific radiologic experience. This could also be an issue for urologists with interest in prostate imaging and fusion biopsy.

We evaluated the role of structured reporting using different PI-RADS versions in a student`s training controlled by in-bore MRI-guided biopsy.

Methods

This retrospective study was approved by the institutional review board and informed consent was obtained from all patients.

Patients and diagnostic MRI

Multiparametric MRI was performed in patients with clinically suspected prostate cancer on a 1.5 Tesla wide bore system (ESPREE, Siemens Healthineers, Erlangen, Germany). For diagnostic MRI, an endorectal coil was used combined to a six-channel body array and an eight-channel spine coil. Our standard protocol fulfilling the requirements of Consensus Meeting on the Standardization of Prostate MRI[9] included T_2-weighted imaging in three planes (slice thickness 3 mm without a gap); the axial in-plane resolution was 0.6 × 0.7 mm. DWI was performed using 4 b-values (0,100, 800, 1400), apparent diffusion coefficient was calculated from b = 0, 100 and 800 s/mm2, the additional b1400 s/mm2 sequence was used for further visual interpretation. Volume interpolated gradient echo sequences with a temporal resolution of 8 s were used for DCE. Gadoteric acid was applied as contrast media in a weight-adapted standard dose (0.1 mmol/kg body weight) with an injection rate of 3 ml s^{-1}. Post-processing of DCE datasets was carried out on a Syngo Multimodality Workplace with Tissue 4D (Siemens, Erlangen, Germany) with colour-coded Ktrans maps and curve evaluation according to Tofts pharmacokinetic model. Lesions scored PIRADS category three or above in multiparametric MRI were assigned to biopsy in a consensus-based decision by radiologist and urologist. After information, 180 consecutive patients (mean age 64.9 years, range 39–82 years) gave their consent. Fifty-eight patients had an ultrasound-guided biopsy without malignant findings before. Median prostate-specific antigen PSA was 7.2 ng ml^{-1} (range 0.4–122.7 ng ml^{-1}, IQR 4.9) with a median PSA density of 0.21 ng/ccm (range 0.02–2.32 ng/ccm, IQR 0.14).

In-bore MR-guided biopsy and lesion marking

Biopsy was performed after accurate biopsy planning in a second session with the patient in supine position lying on a biopsy positioning device (Invivo, Schwerin, Germany) in the same 1.5 Tesla scanner. An adjustable needle guide was inserted transrectally under control of sagittal and coronal fast imaging. Biopsy was performed by an MR compatible fully automatic biopsy gun (18G). Each procedure was documented by fast T_2-weighted imaging in axial and sagittal plane with the needle inside the lesion. It was intended to obtain a minimum of two cores per index lesion. Where appropriate biopsies were repeated, so 1–7 cores (mean 3.8) cores were taken per patient.

For image analysis, the MR scans were reviewed by the radiologist who performed the biopsy. The documented needle position was marked in the axial T2w images of the preceding diagnostic MRI in a synopsis of biopsy and diagnostic scans (Figures 1 and 2). According to the core length of 17 mm, the markings were set in four or five consecutive 3 mm slices

Image evaluation

The medical student in her fifth year did not have prior specific radiological knowledge nor practical experience in MR image evaluation. The student was instructed using the published criteria of PI-RADS V1 in a training less than two hours. The instruction was performed by a radiologist with more than 10 years of experience in prostate imaging and biopsy. The instruction went step by step through the image criteria for T_2-weighted imaging, diffusion-weighted imaging, and DCE based on published sample images for PI-RADS V1.[28] These sample images were available for the student during the evaluation

Figure 2. Same patient as Figure 1; diagnostic imaging; a: T2-weighted axial imaging with retrospectively set markings of two different biopsy cores (arrows), b: contrast-enhanced imaging with diffuse enhancement because of associated prostatitis; c: apparent diffusion coefficient map, d: diffusion-weighted imaging with b-value of 1400; right peripheral lesion: Gleason 3 + 3 (prognostic grade group according to the International Society of Urological Pathology ISUP 1)

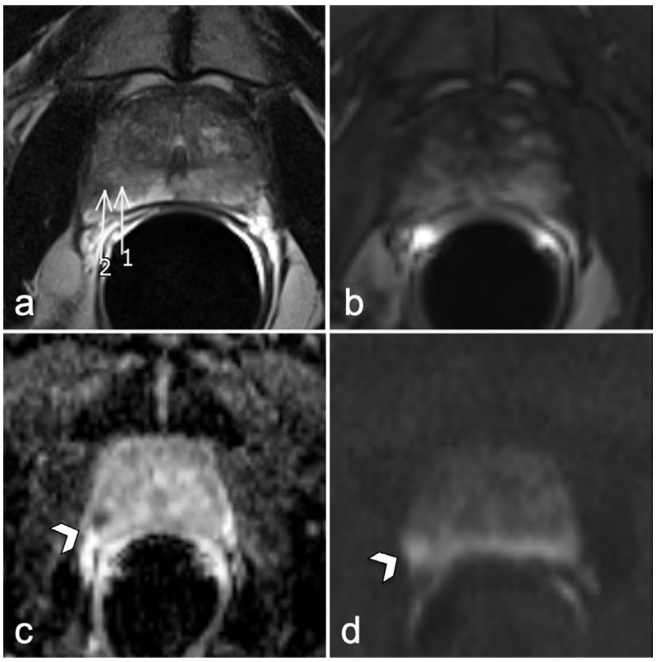


Figure 3. Anterior carcinoma; patient 70 years; prostate-specific antigen PSA 29 ng ml^{-1}; prostate volume 41 ml, PSA density 0.71 ng/ml^2; a–d = diagnostic imaging; a: T2-weighted axial imaging with retrospectively set markings of three different biopsy cores (arrows), b: diffusion-weighted imaging with b-value of 1400; c: apparent diffusion coefficient map, d: contrast-enhanced imaging; *e* + f=biopsy imaging; axial and sagittal documentation of needle position No.2; anterior lesion: Gleason 4 + 4 (prognostic grade group according to the International Society of Urological Pathology ISUP IV)

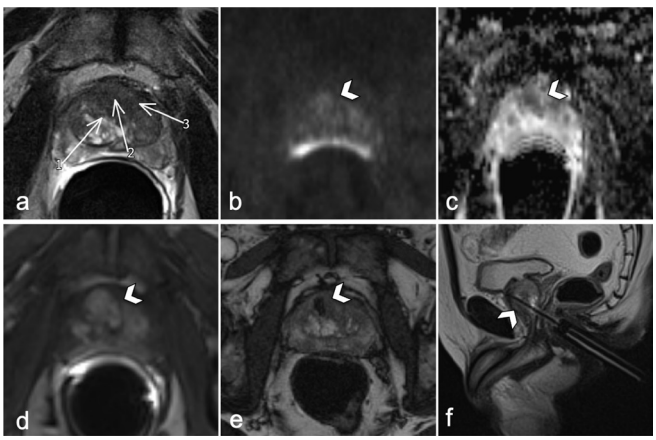
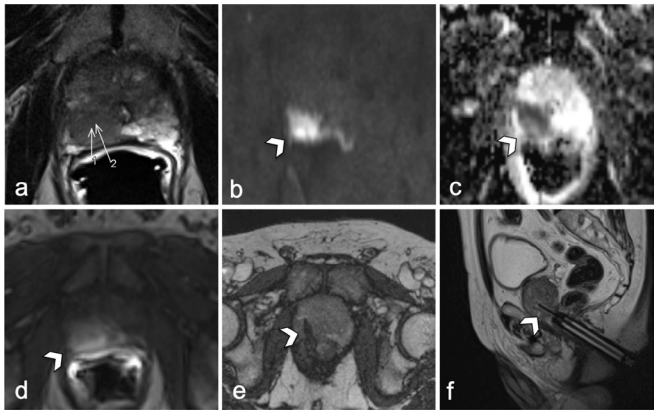


Figure 4. Prostatitis; patient 77 years; prostate-specific antigen PSA 21.9 ng ml^{-1}; prostate volume 85 ml, PSA density 0.26 ng/ml^2; a-d = diagnostic imaging; a: T2-weighted axial imaging with retrospectively set markings of two different biopsy cores (arrows), b: diffusion-weighted imaging with b-value of 1400; c: apparent diffusion coefficient map, d: contrast-enhanced imaging; *e* + f=biopsy imaging; axial and sagittal documentation of needle position No.1; posterior lesion: acute prostatitis, no carcinoma



process. Multiparametric MRI including the markings of 688 cores obtained were displayed in a picture archiving and communicating system to provide an evaluation of the prostate lesions in multiplanar T_2-weighted imaging, in diffusion-weighted imaging including apparent diffusion coefficient maps and DCE (Figures 2–4). Because of missing data 50 of 688 lesions had to be excluded. The documentation of 638 prostate lesions in 180 patients was rated separately by the student and by the radiologist corresponding to PI-RADS V1. After completion of the PI-RADS V1 evaluation, further training on PI-RADS V2 was based on published sample images by Weinreb et al.[11] All data sets were again evaluated corresponding to V2. During the evaluation process, the pathological results were not available.

According to minor differences between V2 and V2.1, the documented results of V2 were searched for lesions of the transition zone (TZ). TZ lesions with T2w single score of 2 were updated to an overall score of 3 in case of a single DWI score of ≥4.

Data analysis

The results were correlated with the presence or absence of prostate carcinoma in the tissue samples obtained. Diagnostic performances of the student and experienced radiologist were evaluated for different cutpoints including sensitivity, specificity and correct-graded samples. Area under the receiver operating characteristic (ROC) curves (AUCs) was calculated. The agreement between student and radiologist was determined using Cohen's weighted κ analysis for the aggregation of the PI-RADS score in three categories according to the clinical relevance. PI-RADS scores 1 and 2 were considered negative, PI-RADS score 3 was considered equivocal and PI-RADS scores 4 and 5 were considered positive for suspected cancer. The time needed for score assessment of the student for V1 and V2 was measured and compared using the Wilcoxon signed-rank test.

Table 1. Rating by the student and by the radiologist according to PI-RADS version 1 (V1) and PI-RADS version 2 (V2)

PI-RADS	Student V1 n (%)	Student V2 (%) n (%)	Radiologist V1 n (%)	Radiologist V2 n (%)
1	3 (<1)	29 (5)	28 (4)	41 (6)
2	150 (24)	263 (41)	116 (18)	255 (40)
3	247 (39)	129 (20)	243 (38)	108 (17)
4	173 (27)	180 (28)	132 (21)	161 (25)
5	65 (10)	37 (6)	119 (19)	73 (11)

The Shapiro-Wilk test was used to assess the normality of the distribution of continuous variables.

Results

The majority of the 638 samples obtained yielded benign pathology (n = 443; 69 %). Prostate cancer was detected in 195 samples in 88 patients. Of these malignant samples, 157 lesions were located in the peripheral zone and 38 lesions in the transition zone. The mean size of the related suspect area in MRI was 11.1 mm (SD 4.5, range 4–22 mm). The prognostic grade group according to the International Society of Urological Pathology (ISUP) 1, 2, 3, 4 and 5 was determined in 67 (34 %), 46 (24 %), 47 (24 %), 27 (14 %) and 8 (4 %) of the malignant cores obtained.

The results of the PI-RADS score according to V1 and V2 rating by the student and by the radiologist are given in Table 1. The distribution of the scores in Figure 5 shows that in both ratings the equivocal score PI-RADS three was less frequent using V2 compared to V1 while the non-suspect score PI-RADS two was more frequent.

The ROC analysis is based on sensitivity, specificity and correct graded samples for different cutpoints (Tables 2 and 3). The student achieved an area under curve (AUC) of 0.782 (95% confidence interval: [0.745; 0.819]) using V1 and 0.788 (95%-confidence interval: [0.750; 0.826]) using V2. The radiologist achieved an AUC of 0.841 (95% confidence interval: [0.806; 0.876]) using V1 and 0.833 (95% confidence interval: [0.799; 0.867]) using V2. The difference between student and radiologist was significant for V1 (p = 0.0002) and V2 (p = 0.002).

The agreement between student and radiologist showed a Cohen's weighted κ coefficient of 0.495 (95% confidence interval: [0.443;0.546]) for V1 and 0.518 for V2 (95% confidence interval: [0.464;0.573]).

The median time needed for score assessment of the student was 4:34 min for V1 and 2:00 min for V2 (p < 0.001).

For re-evaluation of V2.1, 236 lesions of the transition zone were identified. TZ lesions were updated from a V2 score of 2 to a V2.1 score of 3 in case of a DWI category of ≥4. The student´s rating was updated for five lesions (2% of all TZ lesions), none of these five lesions showed cancer. The radiologist's rating was updated for 13 lesions (6% of all TZ lesions); four of these 13 lesions showed cancer. Overall, 1.4% of all ratings had to be changed from V2 to V2.1.

DISCUSSION

Multiparametric prostate MRI represents a complex examination with a high clinical impact. The decision whether to perform a prostate biopsy or not is increasingly based on MRI results besides clinical factors. Inter-reader variability limits this strength of MRI. The clinical management of patients can be altered in a high percentage when a second opinion is considered.[29] Structured reporting in education using PI-RADS may be appropriate to improve diagnostic accuracy.

In both ratings of the student and the experienced radiologist, the equivocal score PI-RADS three was less frequent in favour of PI-RADS two using V2 compared to V1.

Figure 5. Ratings by the student and by the radiologist V1 versus V2 (n = 638 samples). PI-RADS score three is less frequent and PI-RADS score two is more frequent using V2.

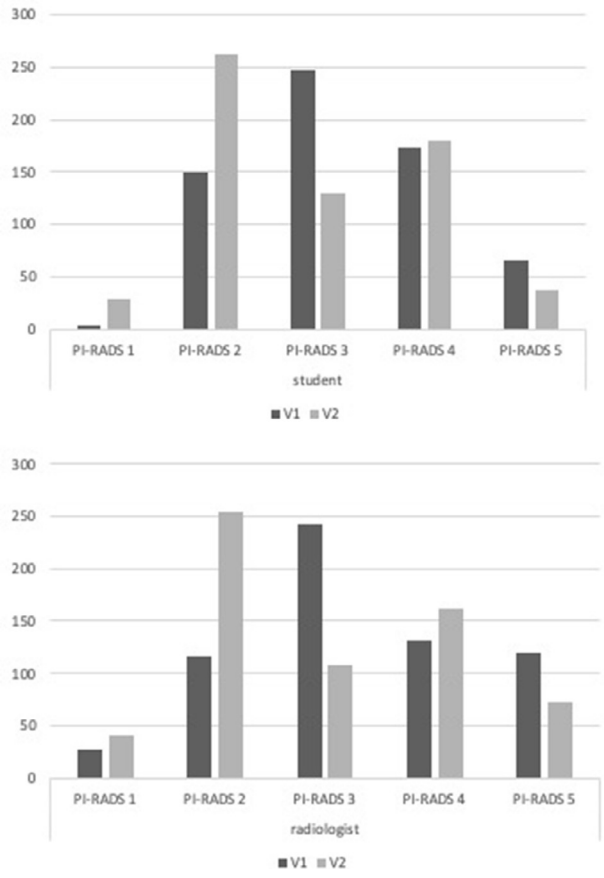


Table 2. Rating by the student "benign" or "malignant" - V1 versus V2; sensitivity, specificity and correct-graded samples for different cutpoints

Cutpoint	Sensitivity (%) V1/V2	Specificity (%) V1/V2	Correct grades samples (%) V1/V2
>=1	100.00/100.00	0.00/0.00	29.77/30.22
>=2	99.48/98.45	0.66/5.58	30.08/33.64
>=3	94.27/83.51	32.67/57.14	51.01/65.11
>=4	70.83/73.20	77.26/81.25	75.35/78.82
>=5	25.00/14.95	96.47/97.77	75.19/72.74
>5	0.00/0.00	100.00/100.00	70.23/69.78

For clinical use, this might help in decision-making, whether a biopsy is indicated or not.

Greer et al differentiated agreement of radiologists with different levels of experience for V2 categories 3 and 4 and found slight agreement for V2 category 3 but substantial agreement for V2 category 4.[30]

Diagnostic performance of different versions of PI-RADS

Many comparative studies, that focused diagnostic performance V1 versus V2 on experienced readers, showed different results. Some studies did not prove any significant difference between V1 and V2.[14–16,18,31,32] Other studies showed lower accuracy of V2.[19,20] Improved diagnostic performance of V2 versus V1 was also described.[21,22]

Reevaluation of V2.1 resulted in minor changes in only 6% of all TZ nodules, yielding additional cancer in less than 1% in our study. This is in line with the results of Byun et al, who reported an upgrade in 4% (eight cases in 201 TZ lesions) with additional cancer in 2% (four cases).[23] Lim et al evaluated 40 upgraded TZ lesions; 27.5% (11/40) showed cancer and 7.5% (3/40) showed significant cancer.[33]

Studies comparing V2 to V2.1 yielded mixed results for diagnostic performance and inter -observer agreement. Wei et al are in favour of V2.1,[25] whereas others did not prove significant benefit for V2.1 versus V2.[23,26,27]

In summary in the literature, there is no unequivocally endorsed version of PI-RADS for experienced readers.

Experience-dependent diagnostic performance

There were minor differences of AUC between V1 and V2 for the student (0.782 and 0.788) and for the radiologist (0.841 and 0.833). The radiologist achieved a significantly higher AUC than the student with V1 and V2. This reflects a better performance of the experienced radiologist as expected. This is also confirmed by other authors. When focused to experience-dependent performance in V2, Greer et al found a higher specificity on a patient basis in readings by highly experienced readers compared to intermediate and low level experienced readers.[30] Pickershill et al described significant differences in accuracy by radiologists of different levels of experience in V1(34). None of these studies evaluated performance of a student with no specific knowledge in radiology.

Experience-dependent inter-observer agreement

The agreement between student and radiologist was moderate using V1 and V2 with a slight advantage for V2 according to Cohen's weighted κ.

This is in line with the results of Tewes et al, who reported fair inter -observer agreement for V1 and moderate agreement with V2 in a study with two readers.[16] Other comparative studies evaluated different readers and found poor agreement of individuals and teams for V1,[34] moderate agreement using V2[30,35] or using both, V1 and V2,[31] good agreement for V1 and V2.[15] Substantial

Table 3. Rating by the radiologist "benign" or "malignant" - V1 versus V2; sensitivity, specificity and correct-graded samples for different cutpoints

Cutpoint	Sensitivity (%) V1/V2	Specificity (%) V1/V2	Correct graded samples (%) V1/V2
>=1	100.00/100.00	0.00/0.00	29.91/30.09
>=2	98.44/98.48	5.78/8.26	33.49/35.41
>=3	94.27/87.88	30.22/59.78	49.38/68.24
>=4	79.69/82.83	78.00/81.30	78.50/81.76
>=5	53.65/28.79	96.44/95.87	83.64/75.68
>5	0.00/0.00	100.00/100.00	70.09/69.91

agreement was described for V1 and V2 evaluation limited to anterior lesions.[14] Krishna et al reported a range from slight (DCE) to substantial (DWI) agreement for V1 but moderate for V2.[20] Substantial to almost perfect agreement for V1 and V2 was found by Polanec et al.[18]

Some studies specified the results for different levels of experience. Greer et al also found moderate agreement of different readers with high, intermediate and low experience using V2.[30] Excellent agreement was found only for index lesions, which was not separately evaluated in our study. Kasel-Seibert et al described an increased inter-observer reliability of both, experienced and unexperienced readers, from moderate in V1 to substantial in V2.[21] Glazer et al also evaluated index lesions using V2, readers had experience from 1 to 11 years. Inter-observer agreement was moderate in the peripheral zone and fair in the transition zone.[36] We did not evaluate inter-observer agreement for separate prostate zones.

Our study design was demanding in some respects. Firstly, a student with no specific knowledge in radiology evaluated multiparametric MRI after a training of less than two hours. Secondly, we compared PI-RADS categories to the presence or absence of prostate carcinoma in each tissue sample obtained in MR-guided biopsy. A patient-based evaluation or a limited evaluation of index lesions could have increased the agreement. Nevertheless, our number of patients is considerably higher than in previous studies mentioned above.

Overall, the student yielded a respectable diagnostic performance and inter -observer agreement using structured reporting.

Time saving

In structured reporting, the efficiency in daily routine should be considered.

V2 emphasizes the weight of the diffusion score for peripheral lesions and the weight of the T2w score in transition zone lesions. Consequently, in some cases, it is not mandatory to evaluate the other parameters in detail leading to savings in time. According to this, simpler evaluation in an abbreviated pathway in V2 versus V1 the time needed for score assessment was significantly less (2:00 min versus 4:34 ($p < 0.001$)).

This is in line with the study of Tewes et al, who also reported a significant lower time required for V2 compared to V1.[16]

Education model

This study aimed to quantify the efficiency of structured reporting combined with simple image-based assistance. Clinical cases were prepared for re-evaluation in an education model. The correlation between a student and an experienced reader and the correlation with histological results was measured to compare different versions of the structured reporting.

An education set and a test set for training radiology on breast imaging according to BI-RADS residents have already been evaluated with good results.[37]

The need for novel ways of training in prostate MRI has been described in a survey among urology residents. Online modules for self-directed learning were endorsed.[38] Artificial intelligence was also proposed to make a virtual experienced reader available during the training.

Barth et al evaluated two different readings of non-specialized radiology residents. For the first read, the residents used the PI-RADS v2.1 document only. A second read was assisted by a browser-based calculator.[39] This calculator presented a reporting pathway comparable to the described pathway in our study; based on the mentioned above simplification of V2 versus V1, not all parameters had to be described in detail in this pathway. The student in our study was similarly assisted by visual examples. Barth et al found a significant time-saving using the calculator versus using the PI-RADS document only without loss of diagnostic accuracy. This observation is in line with the significant time-saving in our study in V2 versus V1. Therefore, a clear reporting pathway should be an essential part of structured reporting and effective training. The browser-based presentation may have a further effect on the time required. Further attempts at improvement could include a self-control system based on histological results for predefined lesions. Such a model should be evaluated with participation of several readers.

An optimal education model yields high acceptance and a high trainee outcome.

Our findings may enhance this approach for prostate image interpretation for students, residents in radiology, or urology. The reliability of reporting becomes quantifiable using PI-RADS structured reporting. V2 seems to be slightly superior to V1 regarding reliability and time efficiency in education.

Limitations

Structured reporting according to PI-RADS was reduced in our study design to evaluation of previously described lesions with probability of cancer risk. There was no need to identify lesions independently. No extra prostatic disease and no incidental findings other than prostate carcinoma were considered. However, the results of our study encourage a training model using predefined lesions. A further step could be the autonomous identification of an index lesion.

The allocation of biopsy samples to lesions in MRI (Figures 1 and 2) seems reliable, as biopsy was performed in-bore under image control. Nevertheless, this procedure could be afflicted with errors, as there was no postoperative whole mount histopathological correlation available.

Although during the retrospective evaluation process, the pathological results were not available, the radiologist, who performed the diagnostic and the biopsy, could have been influenced by information from the clinical process. However, in consideration of the high number of lesions, this restriction seems negligible.

The number of lesions was high leading to increasing experience of the student. Training effects of the student between the reading sessions have not been evaluated. This might have contributed to a better inter -observer agreement using V2.

Reevaluation of V2.1 was limited to upgrade of TZ lesions category two in T2w in case of DWI category ≥4 in our study. The modifications in v2.1 also address interpretation of TZ nodules in T2w, this was not evaluated in our study. Furthermore, interpretation of PZ lesions is modified in V2.1. Nevertheless, higher sensitivity for V2.1 in PZ lesions was shown in a current paper only among experienced readers with less specificity at the same time.[27]

CONCLUSION

In our large study population, the capacity of prostate cancer detection using PI-RADS V1 and V2 was dependent on the reader's experience, but with good diagnostic performance even for a student with limited expert radiological knowledge. The results from the two observers indicate that structured reporting using PI-RADS and, controlled by histopathology, can be a valuable and quantifiable tool in students´ or residents´ education. Herein, V2 was superior to V1 in terms of inter -observer agreement and time efficacy.

DECLARATIONS OF INTEREST

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

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

1. Ganeshan D, Duong P-AT, Probyn L, Lenchik L, McArthur TA, Retrouvey M, et al. Structured reporting in radiology. *Acad Radiol* 2018; **25**: 66–73. doi: https://doi.org/10.1016/j.acra.2017.08.005
2. Cury RC, Abbara S, Achenbach S, Agatston A, Berman DS, Budoff MJ, et al. Coronary Artery Disease - Reporting and Data System (CAD-RADS): An Expert Consensus Document of SCCT, ACR and NASCI: Endorsed by the ACC. *JACC Cardiovasc Imaging* 2016; **9**: 1099–113. doi: https://doi.org/10.1016/j.jcmg.2016.05.005
3. Pinto Dos Santos D, Hempel J-M, Mildenberger P, Klöckner R, Persigehl T. Structured reporting in clinical routine. *Rofo* 2019; **191**: 33–9. doi: https://doi.org/10.1055/a-0636-3851
4. Shaish H. Structured prostate MRI reporting: how and why. *Abdom Radiol* 2020; **45**: 3969–73. doi: https://doi.org/10.1007/s00261-020-02720-2
5. Garcia-Reyes K, Passoni NM, Palmeri ML, Kauffman CR, Choudhury KR, Polascik TJ, et al. Detection of prostate cancer with multiparametric MRI (mpMRI): effect of dedicated reader education on accuracy and confidence of index and anterior cancer diagnosis. *Abdom Imaging* 2015; **40**: 134–42. doi: https://doi.org/10.1007/s00261-014-0197-7
6. Nobel JM, Kok EM, Robben SGF. Redefining the structure of structured reporting in radiology. *Insights Imaging* 2020; **11**: 10. doi: https://doi.org/10.1186/s13244-019-0831-6
7. European Society of R. ESR paper on structured reporting in radiology. *Insights Imaging* 2018; **9**: 1–7.
8. Culp MB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. Recent global patterns in prostate cancer incidence and mortality rates. *Eur Urol* 2020; **77**: 38–52. doi: https://doi.org/10.1016/j.eururo.2019.08.005
9. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate Mr guidelines 2012. *Eur Radiol* 2012; **22**: 746–57. doi: https://doi.org/10.1007/s00330-011-2377-y
10. Rosenkrantz AB, Oto A, Turkbey B, Westphalen AC. Prostate imaging reporting and data system (PI-RADS), version 2: a critical look. *AJR Am J Roentgenol* 2016; **206**: 1179–83. doi: https://doi.org/10.2214/AJR.15.15765
11. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* 2016; **69**: 16–40. doi: https://doi.org/10.1016/j.eururo.2015.08.052
12. 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 prostate imaging reporting and data system version 2. *Eur Urol* 2019; **76**: 340–51. doi: https://doi.org/10.1016/j.eururo.2019.02.033
13. Schaudinn A, Gawlitza J, Mucha S, Linder N, Franz T, Horn L-C, et al. Comparison of PI-RADS V1 and V2 for multiparametric MRI detection of prostate cancer with whole-mount histological workup as reference standard. *Eur J Radiol* 2019; **116**: 180–5. doi: https://doi.org/10.1016/j.ejrad.2019.04.012
14. Hoffmann R, Logan C, O'Callaghan M, Gormly K, Chan K, Foreman D. Does the prostate Imaging-Reporting and data system (PI-RADS) version 2 improve accuracy in reporting anterior lesions on multiparametric magnetic resonance imaging (mpMRI)? *Int Urol Nephrol* 2018; **50**: 13–19. doi: https://doi.org/10.1007/s11255-017-1753-1
15. Becker AS, Cornelius A, Reiner CS, Stocker D, Ulbrich EJ, Barth BK, 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* 2017; **94**: 58–63. doi: https://doi.org/10.1016/j.ejrad.2017.07.016
16. Tewes S, Mokov N, Hartung D, Schick V, Peters I, Schedl P, et al. Standardized reporting of prostate MRI: comparison of the prostate imaging reporting and data system (PI-RADS) version 1 and version 2. *PLoS One* 2016; **11**: e0162879. doi: https://doi.org/10.1371/journal.pone.0162879
17. Tan N, Lin W-C, Khoshnoodi P, Asvadi NH, Yoshida J, Margolis DJA, et al. In-Bore 3-T MR-guided transrectal targeted prostate biopsy: prostate imaging reporting and data system version 2–based diagnostic performance for detection of prostate cancer. *Radiology* 2017; **283**: 130–9. doi: https://doi.org/10.1148/radiol.2016152827
18. Polanec S, Helbich TH, Bickel H, Pinker-Domenig K, Georg D, Shariat SF, et al. Head-To-Head comparison of PI-RADS V2 and PI-RADS V1. *Eur J Radiol* 2016; **85**: 1125–31. doi: https://doi.org/10.1016/j.ejrad.2016.03.025
19. Auer T, Edlinger M, Bektic J, Nagele U, Herrmann T, Schäfer G, et al. Performance of PI-RADS version 1 versus version 2

regarding the relation with histopathological results. *World J Urol* 2017; **35**: 687–93. doi: https://doi.org/10.1007/s00345-016-1920-5
20. Krishna S, McInnes M, Lim C, Lim R, Hakim SW, Flood TA, et al. Comparison of prostate imaging reporting and data system versions 1 and 2 for the detection of peripheral zone Gleason score 3 + 4 = 7 cancers. *AJR Am J Roentgenol* 2017; **209**: W365–73. doi: https://doi.org/10.2214/AJR.17.17964
21. Kasel-Seibert M, Lehmann T, Aschenbach R, Guettler FV, Abubrig M, Grimm M-O, et al. Assessment of PI-RADS V2 for the detection of prostate cancer. *Eur J Radiol* 2016; **85**: 726–31. doi: https://doi.org/10.1016/j.ejrad.2016.01.011
22. Feng Z-Y, Wang L, Min X-D, Wang S-G, Wang G-P, Cai J. Prostate cancer detection with multiparametric magnetic resonance imaging: prostate imaging reporting and data system version 1 versus version 2. *Chin Med J* 2016; **129**: 2451–9. doi: https://doi.org/10.4103/0366-6999.191771
23. Byun J, Park KJ, Kim M-H, Kim JK. Direct comparison of PI-RADS version 2 and 2.1 in transition zone lesions for detection of prostate cancer: preliminary experience. *J Magn Reson Imaging* 2020; **52**: 577–86. doi: https://doi.org/10.1002/jmri.27080
24. Lim CS, Abreu-Gomez J, Carrion I, Schieda N. Prevalence of prostate cancer in PI-RADS version 2.1 transition zone atypical nodules upgraded by abnormal DWI: correlation with MRI-Directed TRUS-Guided targeted biopsy. *American Journal of Roentgenology* 2021; **216**: 683–90. doi: https://doi.org/10.2214/AJR.20.23932
25. Wei CG, Zhang YY, Pan P, Chen T, HC Y, Dai GC. Diagnostic accuracy and inter-observer agreement of PI-RADS version 2 and version 2.1 for the detection of transition zone prostate cancers. *AJR American journal of roentgenology* 2020;.
26. Wang Z, Zhao W, Shen J, Jiang Z, Yang S, Tan S, et al. PI-RADS version 2.1 scoring system is superior in detecting transition zone prostate cancer: a diagnostic study. *Abdominal Radiology* 2020; **45**: 4142–9. doi: https://doi.org/10.1007/s00261-020-02724-y
27. 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**: 141-1519(Online ahead of print). doi: https://doi.org/10.2214/AJR.20.24199
28. Röthke M, Blondin D, Schlemmer H-P, Franiel T. PI-RADS classification: structured reporting for MRI of the prostate. *Rofo* 2013; **185**: 253–61. doi: https://doi.org/10.1055/s-0032-1330270
29. Luzzago S, Petralia G, Musi G, Catellani M, Alessi S, Di Trapani E, et al. Multiparametric magnetic resonance imaging second opinion may reduce the number of unnecessary prostate biopsies: time to improve radiologists' training program? *Clin Genitourin Cancer* 2019; **17**: 88–96. doi: https://doi.org/10.1016/j.clgc.2018.10.006
30. Greer MD, Shih JH, Lay N, Barrett T, Bittencourt L, Borofsky S, et al. Interreader variability of prostate imaging reporting and data system version 2 in detecting and assessing prostate cancer lesions at prostate MRI. *AJR Am J Roentgenol* 2019; **27**(Mar): 1197–205. doi: https://doi.org/10.2214/AJR.18.20536
31. Schaudinn A, Gawlitza J, Mucha S, Linder N, Franz T, Horn L-C, Alexander Schaudinn JG, Kahn T, et al. Comparison of PI-RADS V1 and V2 for multiparametric MRI detection of prostate cancer with whole-mount histological workup as reference standard. *Eur J Radiol* 2019; **116**(Jul): 180–5. doi: https://doi.org/10.1016/j.ejrad.2019.04.012
32. Tan N, Lin W-C, Khoshnoodi P, Asvadi NH, Yoshida J, Margolis DJA, et al. In-Bore 3-T MR-guided transrectal targeted prostate biopsy: prostate imaging reporting and data system version 2-based diagnostic performance for detection of prostate cancer. *Radiology* 2017; **283**: 130-139. doi: https://doi.org/10.1148/radiol.2016152827
33. Christopher S Lim JA-G, Ivan Carrion, Nicola Schieda. Prevalence of prostate cancer in PI-RADS version 2.1 transition zone 'atypical nodules' upgraded by abnormal diffusion weighted imaging: correlation with MRI-directed TRUS-guided targeted biopsy. *AJR American journal of roentgenology* 2020;.
34. Pickersgill NA, Vetter JM, Andriole GL, Shetty AS, Fowler KJ, Mintz AJ, et al. Accuracy and variability of prostate multiparametric magnetic resonance imaging interpretation using the prostate imaging reporting and data system: a blinded comparison of radiologists. *Eur Urol Focus* 2020; **6**: 267-272. doi: https://doi.org/10.1016/j.euf.2018.10.008
35. Zhao C, Gao G, Fang D, Li F, Yang X, Wang H, et al. The efficiency of multiparametric magnetic resonance imaging (mpMRI) using PI-RADS version 2 in the diagnosis of clinically significant prostate cancer. *Clin Imaging* 2016; **40**: 885–8. doi: https://doi.org/10.1016/j.clinimag.2016.04.010
36. Glazer DI, Mayo-Smith WW, Sainani NI, Sadow CA, Vangel MG, Tempany CM, et al. Interreader agreement of prostate imaging reporting and data system version 2 using an In-Bore MRI-guided prostate biopsy cohort: a single institution's initial experience. *AJR Am J Roentgenol* 2017; **209**: W145–51. doi: https://doi.org/10.2214/AJR.16.17551
37. Yoon JH, Lee HS, Kim YM, Youk JH, Kim SH, Jeong SH, et al. Effect of training on ultrasonography (US) BI-RADS features for radiology residents: a multicenter study comparing performances after training. *Eur Radiol* 2019; **29**: 4468–76. doi: https://doi.org/10.1007/s00330-018-5934-9
38. Rodrigues C, Visram K, Sedghi A, Mousavi P, Siemens DR. Attitudes and experience of urology trainees in interpreting prostate magnetic resonance imaging. *Can Urol Assoc J* 2021; **15**: E293–8. doi: https://doi.org/10.5489/cuaj.6614
39. Barth BK, Martini K, Skawran SM, Schmid FA, Rupp NJ, Zuber L, et al. Value of an online PI-RADS v2.1 score calculator for assessment of prostate MRI. *Eur J Radiol Open* 2021; **8**: 100332. doi: https://doi.org/10.1016/j.ejro.2021.100332