A Prospective Accuracy Study of Prostate Imaging Reporting and Data System Version 2 on Multiparametric Magnetic Resonance Imaging in Detecting Clinically Significant Prostate Cancer With Whole-mount Pathology

Gianluca Giannarini, Rossano Girometti, Alessandro Crestani, Marta Rossanese, Mattia Calandriello, Lorenzo Cereser, Sandra Bednarova, Claudio Battistella, Stefano Sioletic, Chiara Zuiani, Claudio Valotto, and Vincenzo Ficarra

OBJECTIVE To assess the accuracy of Prostate Imaging Reporting and Data System version 2 (PI-RADS v2) in detecting clinically significant prostate cancer (csPCa) on multiparametric magnetic resonance imaging (mpMRI) using whole-mount sections after radical prostatectomy (RP) as reference standard.

METHODS Forty-eight patients undergoing mpMRI before RP were prospectively enrolled. Two experienced radiologists independently scored and mapped imaging findings according to PI-RADS v2. One experienced uropathologist mapped cancers detected on whole-mount sections using the PI-RADS v2 sector scheme. Per-lesion and per-patient analyses were run. Primary outcomes were sensitivity and false discovery rate (FDR) in detecting csPCa using PI-RADS v2 score ≥3 and ≥4 as thresholds. Secondary outcome was inter-reader agreement.

RESULTS On the per-lesion analysis, sensitivity and FDR at the PI-RADS v2 threshold score ≥3 were 0.75 and 0.17 for Reader 1, and 0.67 and 0.13 for Reader 2, respectively. At the PI-RADS v2 threshold score ≥4, sensitivity was slightly lower, and FDR nearly halved for both readers. On the per-patient analysis, sensitivity for csPCa at the PI-RADS v2 threshold score ≥3 was 0.85 for Reader 1, and 0.78 for Reader 2. At the PI-RADS v2 threshold score ≥4, sensitivity was slightly lower for both readers. Inter-reader agreement was substantial (k 0.72 and 0.65 for PI-RADS v2 threshold score ≥3 and ≥4, respectively).

CONCLUSION In our prospective study with pathology after RP as standard of reference, PI-RADS v2 showed good sensitivity in detecting csPCa on mpMRI with substantial agreement between 2 experienced readers. Threshold score ≥4 had lower FDR. UROLOGY 123: 191–197, 2019. © 2018 Published by Elsevier Inc.

Accumulating evidence suggests that multiparametric magnetic resonance imaging (mpMRI) may improve the detection of clinically significant prostate cancer (csPCa).[1] This modality, however, requires great expertise and it is still limited by significant inter-reader variability.[2] With its increasing dissemination, continuous efforts should be taken to ameliorate image interpretation and reporting across readers and centers.

One major step toward standardization of mpMRI has been the introduction of Prostate Imaging Reporting and Data System (PI-RADS) in 2012.[3] In this system, a suspicion score for the presence of csPCa was assigned on a 1-5 scale on each of the 3 mpMRI sequences, but no standard criteria to calculate an overall PI-RADS score were established. A systematic review with meta-analysis showed that PI-RADS had good diagnostic accuracy in detecting prostate cancer (PCa), with a pooled sensitivity and

From the Urology Unit, Academic Medical Centre "Santa Maria della Misericordia", Udine, Italy; the Department of Medicine, Institute of Radiology, University of Udine, Udine, Italy; the Department of Human and Paediatric Pathology "Gaetano Barresi", Urologic Section, University of Messina, Messina, Italy; the Department of Medicine, University of Udine, Udine, Italy; and the Pathology Unit, Academic Medical Centre "Santa Maria della Misericordia", Udine, Italy

Address correspondence to: Vincenzo Ficarra, M.D, F.E.B.U., Department of Human and Paediatric Pathology "Gaetano Barresi", Urologic Section, University of Messina, via Consolare Valeria 1, Messina 98124 Italy. E-mail: vficarra@unime.it

Submitted: June 6, 2018, accepted (with revisions): July 23, 2018

specificity of 0.78 (95% confidence intervals [CIs] 0.70-0.84) and 0.79 (95% CIs 0.68-0.86), respectively. However, no recommendations regarding the best threshold for PI-RADS to use in clinical practice could be provided in this review because of the large expected between-study heterogeneity.[4]

In order to improve accuracy, a novel version (ie, PI-RADS v2) was released in late 2015 with precise recommendations such as integrating mpMRI scores according to prostate zonal anatomy into a single level of suspicion of csPCa on a 5-point scale.[5] In a recent systematic review with meta-analysis, PI-RADS v2 showed a significantly higher pooled sensitivity (0.95 [95% CIs 0.85-0.98] vs 0.88 [95% CIs 0.80-0.93]), but a similar specificity (0.73 [95% CIs 0.47-0.89] vs 0.75 [95% CIs 0.36-0.94], for PCa detection compared to original PI-RADS.[6]

Intrinsic limitations to both analyses were paucity of well-powered prospective studies, heterogeneity in the definition of csPCa, and almost exclusive correlation with histology at prostate biopsy.

The aim of our study was to prospectively validate the PI-RADS v2 on 3.0 Tesla mpMRI in detecting csPCa using whole-mount pathology after radical prostatectomy (RP) as the standard of reference.

METHODS

Patients

A prospectively maintained database for RP patients has been active at the Urology Unit of the Academic Medical Centre in Udine since July 1, 2013.

This prospective single-center diagnostic study was performed as a part of a larger prospective institutional review board-approved investigation of the role of 3.0T mpMRI on PCa management, and was conducted in accordance with the Declaration of Helsinki. Starting from May 1, 2016, 3.0T prostate mpMRI was introduced as a routine examination before RP in the frame of a quality improvement program running at our institution. PI-RADS v2 was routinely adopted for image reporting since that time. With the aim of validating PI-RADS v2 for the detection of csPCa, all consecutive patients scheduled for RP from that time to February 28, 2017 were included in the present study. All patients were diagnosed with systematic 12-core biopsy. Exclusion criteria were the previous treatments for PCa, a time interval between prostate biopsy and mpMRI <6 weeks in line with PI-RADS v2 recommendations,[5] and a time interval between mpMRI and RP ≥4 months. Written informed consent was obtained by all participants.

MRI Technique and Image Analysis

All mpMRI examinations were performed on a 3.0T system (Achieva, Philips Medical System, Best, the Netherlands), using a 32-channel phased-array surface coil and no endorectal coil. Details of mpMRI protocol are reported in Supplementary Table 1.

This was set as a multireader study. Reading sessions were organized on a dedicated workstation (Olea Sphere, Olea Medical, La Ciotat, France), where image sets were anonymized with respect to all patient identifiers. For the purpose of this analysis, 2 radiologists (GR and LC with 8 and 6 years of experience in interpreting prostate mpMRI), who were blinded to clinical data and final pathology, independently analyzed all images. Both readers completed a 3-month training period using PI-RADS v2 in the clinical routine before study initiation. Lesions identified by the readers were scored and mapped on the 36-sector scheme according to PI-RADS v2 criteria.[5] Absence of prostatic lesions was categorized as PI-RADS v2 score 1. Parameters recorded for each lesion were: side (right, left, and bilateral), zone (peripheral zone also including central zone, transition zone also including anterior fibromuscular stroma), level (base, midgland, and apex), and PI-RADS v2 score. After completion of image analysis, each anonymized reader sheet was stored in a secured local server for the final analysis.

Pathology Analysis

All patients underwent retropubic RP. One pathologist with a 10-year experience in genitourinary pathology, who was blind to clinical and imaging data, processed RP specimens using whole-mount protocol according to International Society of Urological Pathology (ISUP) standards.[7] In detail, specimens were fixed in 4% buffered formaldehyde for 48 hours at room temperature. The apical 3-8 mm of the prostate was amputated and longitudinally sliced, and the remaining gland was serially sectioned from the apex to the base in 4-6 mm axial slices perpendicular to the urethra. Slices were submitted for paraffin-embedded whole-mounts from which 4-5 μm slides were stained with hematoxylin and eosin. To allow for an accurate orientation and to keep a relatively equal space between the histology sections of each specimen, each slice of the prostate was accurately cut and orientated. All cancer foci were outlined with a marker on the glass slide. Parameters recorded for each cancer were: maximum diameter, ISUP grade[8], and extraprostatic extension or seminal vesicle invasion if present. Since a diameter of approximately 1 cm corresponds to a spherical volume of 0.5 cc, which is the currently accepted threshold for csPCa,[9] and in line with previous studies,[10] we analyzed tumor maximum diameter as a surrogate of volume. We defined csPCa any tumor with diameter ≥1 cm or ISUP grade ≥2 or extraprostatic extension and/or seminal vesicle invasion. Pathological extent of the primary tumor was assigned according to the 2010 Tumor-Node-Metastasis staging system. Finally, all cancers were mapped in a bidimensional illustration of the prostate sections corresponding to the analyzed whole-mount slides with the purpose to create a reliable 3-dimensional model of the identified tumors. The same sector division as in the PI-RADS v2 scheme was adopted.

Matching of Imaging and Pathology Findings

An independent radiologist, who had 4 years of experience in interpreting prostate mpMRI and completed the same training period with PI-RADS v2 as the 2 study readers, analyzed the 2 maps for each patient together with the uropathologist in order to match each lesion identified on mpMRI to the corresponding findings on definitive histology. Similarly to a previous study,[11] whole-mount slides with the greatest cross-section of the specific lesion were used for concordance analysis of location on axial T2-weighted images. Matching was performed for both PCa and csPCa. In the per-lesion analysis, true positivity of mpMRI lesions was considered if there was exact sector correspondence or a discrepancy between the mpMRI lesion and the pathologic lesion of up to 1 sector in any direction, provided that side and zone were unaltered.

Study Outcomes and Statistical Analyses

Primary outcome was sensitivity of PI-RADS v2 in detecting csPCa using score 3 and 4 as thresholds for test positivity. Both per-lesion and per-patient analyses were carried out, and 2 × 2 contingency tables were used to calculate all accuracy measures as reported in Figure 1. Specificity was deliberately omitted, because mpMRI readers tend to report highly suspicious lesions only and avoid over-reporting additional "negative" lesions or sectors. Instead, false discovery rate (FDR), defined as the ratio between the number of false positive lesions and total number of positive lesions on mpMRI, was calculated as a further measure of accuracy.[12] Secondary outcome was an inter-reader agreement.

Based on literature data, a sample of 46 patients with at least 1 tumor lesion on definitive pathology after RP allowed to estimate 0.7 sensitivity in detecting csPCa with 95% CIs and a 0.14 precision.[13] Based on in-house data, an approximate 0.1 rate of suboptimal mpMRI quality was expected, thus, a total of 53 patients were enrolled.

Continuous variables were reported as mean ± standard deviation or median and interquartile range (IQR) as appropriate. Categorical variables were reported as frequency and percentages. McNemar test was used to compare sensitivity values for csPCa, both between the 2 readers and between the 2 PI-RADS v2 thresholds. Inter-reader agreement was assessed with Cohen's weighted-kappa coefficient and corresponding 95% CIs calculated with bootstrap method.[14] All clinical records were inserted in a dedicated database and data were analyzed using Stata and/or SE 14.1 (StataCorp LP, College Station, TX) software. Reporting of this study was in accordance with the updated Standards for Reporting of Diagnostic Accuracy Studies (Supplementary Table 2).[15]

RESULTS

Of the 53 enrolled patients, 5 with mpMRI examinations of suboptimal quality due to severe gas artefacts were excluded. The final study population comprised 48 patients (Supplementary Figure 1). Mean (standard deviation) age was 65.8 (6.5) years. Median (IQR) preoperative prostate-specific antigen level was 7.2 (5.2-10) ng/mL. Median (IQR) time from prostate biopsy to mpMRI and from mpMRI to RP was 7 (7-7.8) weeks and 2.9 (1.9-3.9) weeks, respectively.

Of the 48 patients, 41 (85%) had csPCa. A total of 71 tumor foci were detected on final histology, and 51 (71.8%) were clinically significant cancers. One focus was detected in 25 (52%) patients, and 2 foci were detected in 23 (48%) patients (Table 1).

Per-lesion analysis	The mpMRI finding was a csPCa on definitive pathology	The mpMRI finding was not a csPCa on definitive pathology
mpMRI finding PI-RADS ≥3 (or ≥4)	True positive	False positive
mpMRI finding PI-RADS <3 (or <4)	False negative	True negative

Per-patient analysis	csPCa on definitive pathology	non csPCa on definitive pathology
At least one mpMRI finding PI-RADS ≥3 (or ≥4)	True positive	False positive
mpMRI finding/s PI-RADS <3 (or <4)	False negative	True negative

Figure 1. Criteria used to calculate measures of diagnostic accuracy according to different PI-RADS v2 threshold score. csPCa, clinically significant prostate cancer; mpMRI, multiparametric magnetic resonance imaging; PI-RADS v2, Prostate Imaging Reporting and Data System version 2.

Table 1. Characteristics of the 71 prostate cancer foci detected on whole-mount sections

Variable		Value
ISUP grading group, n (%)	group 1	41 (57.7)
	group 2	13 (18.3)
	group 3	7 (9.9)
	group 4	7 (9.9)
	group 5	3 (4.2)
Extraprostatic extension and/or Seminal vesicle invasion, n (%)		20 (28.2)
Maximum tumor diameter (mm), n (%)	< 10	23 (32.4)
	≥ 10	48 (67.6)
Zone, n (%)	PZ	57 (80.3)
	TZ	11 (15.5)
	PZ + TZ	3 (4.2)

ISUP, International Society of Urological Pathology; PZ, peripheral zone; TZ, transition zone.

Total number of PI-RADS v2 assignments was 80 for Reader 1 and 78 for Reader 2. PI-RADS v2 score 1, 2, 3, 4, and 5 was assigned to 25, 3, 5, 32, and 15 lesions by Reader 1, and to 30, 4, 8, 17, and 19 lesions by Reader 2, respectively. Results of the correspondence between mpMRI and pathology findings for the 2 readers stratified by PI-RADS v2 threshold and presence of csPCa are detailed in Table 2.

On the per-lesion analysis, sensitivity and FDR for csPCa at the PI-RADS v2 threshold score ≥3 were 0.75 and 0.17 for Reader 1, and 0.67 and 0.13 for Reader 2, respectively. At the PI-RADS v2 threshold score ≥4, sensitivity was slightly lower, and FDR nearly halved for both readers. Details are reported in Table 3. Sensitivity values were comparable between the 2 readers and the 2 PI-RADS v2 thresholds, according to McNemar test. Of the false negative lesions at the PI-RADS v2 threshold score ≥4, 16/28 (57.1%) and 17/37 (45.9%) had a diameter < 1 cm, 22/28 (78.5%) and 25/37 (67.5%) had ISUP grading group 1, and 25/28 (89.2%) and 29/37 (78.3%) were intracapsular, for Readers 1 and 2, respectively.

On the per-patient analysis, sensitivity for csPCa at the PI-RADS v2 threshold score ≥3 was 0.85 for Reader 1, and 0.78 for Reader 2. At the PI-RADS v2 threshold score ≥4, sensitivity was slightly lower for both readers (Table 3). No significant difference was observed comparing sensitivity values between the 2 readers and the 2 PI-RADS v2 thresholds with McNemar test.

Inter-reader agreement was substantial. For csPCa, k was 0.72 (95% CIs 0.51-0.79) and 0.65 (95% CIs 0.57-0.88) for PI-RADS v2 threshold score ≥3 and ≥4, respectively.

DISCUSSION

To our knowledge, this is the first prospective study compliant with Standards for Reporting of Diagnostic Accuracy Studies guidelines that validated the PI-RADS v2 on 3.0 Tesla mpMRI without endorectal coil in detecting csPCa on definitive pathology after RP. We found PI-RADS v2 had a good sensitivity in detecting csPCa with substantial agreement between 2 experienced readers. Using PI-RADS ≥4 as threshold for test positivity resulted in lower FDR.

Caution should be exerted when comparing the results of our study with those of the literature for different reasons.

Sensitivity values at both PI-RADS v2 thresholds in our study were lower than pooled values reported in the previously quoted meta-analysis (ie, 0.96 for threshold score ≥3 and 0.89 for threshold score ≥4).[6] However, most included studies had histology at biopsy as the standard of reference, and the only 5 studies with definitive pathology after RP as the sole standard of reference were

Table 2. Classification of mpMRI findings for the 2 readers stratified by PI-RADS v2 threshold score according to pathology results as defined in Figure 1

Analysis	Outcome	PI-RADS v2 Threshold	Reader	TP, n	FN, n	FP, n	TN, n
Per-lesion	Any PCa	Score ≥ 3	1	44	27	8	1
			2	39	32	5	2
		Score ≥ 4	1	43	28	4	5
			2	34	37	2	5
	csPCa	Score ≥ 3	1	38	13	8	1
			2	34	17	5	2
		Score ≥ 4	1	37	14	4	5
			2	31	20	2	5
Per-patient	Any PCa	Score ≥ 3	1	38	10	–	–
			2	35	13	–	–
		Score ≥ 4	1	37	11	–	–
			2	31	17	–	–
	csPCa	Score ≥ 3	1	35	6	–	–
			2	32	9	–	–
		Score ≥ 4	1	34	7	–	–
			2	30	11	–	–

csPCa, clinically significant prostate cancer; FN, false negative; FP, false positive; mpMRI, multiparametric magnetic resonance imaging; PCa, prostate cancer; PI-RADS v2, Prostate Imaging Reporting and Data System version 2; TN, true negative; TP, true positive.

Table 3. Per-lesion and per-patients analysis of diagnostic performance of PI-RADS v2 in detecting any PCa and csPCa stratified by threshold score

Measure	Analysis	Outcome	Reader	Values (95% CIs)	
				PI-RADS v2 Threshold ≥ 3	PI-RADS v2 Threshold ≥ 4
Sensitivity	Per-lesion	Any PCa	Reader 1	0.62 (0.50–0.73)	0.61 (0.48 – 0.72)
			Reader 2	0.55 (0.43 – 0.67)	0.48 (0.36 – 0.60)
		csPCa	Reader 1	0.75 (0.60 – 0.86)	0.73 (0.58 – 0.84)
			Reader 2	0.67 (0.52 – 0.79)	0.61 (0.46 – 0.74)
	Per-patient	Any PCa	Reader 1	0.79 (0.65 – 0.90)	0.77 (0.63 – 0.88)
			Reader 2	0.73 (0.58 – 0.85)	0.65 (0.49 – 0.78)
		csPCa	Reader 1	0.85 (0.71 – 0.94)	0.83 (0.68 – 0.93)
			Reader 2	0.78 (0.62 – 0.89)	0.73 (0.57 – 0.86)
FDR	Per-lesion	Any PCa	Reader 1	0.15 (0.07 – 0.28)	0.09 (0.02 – 0.20)
			Reader 2	0.11 (0.04 – 0.25)	0.06 (0.01 – 0.19)
		csPCa	Reader 1	0.17 (0.08 – 0.31)	0.10 (0.03 – 0.23)
			Reader 2	0.13 (0.04 – 0.27)	0.06 (0.01 – 0.20)

CIs, confidence intervals; FDR, false discovery rate.

retrospective. It is known that studies using pathology from targeted prostate biopsy have the potential bias of artificially inflating measures of diagnostic accuracy because only pre-selected highly suspicious lesions are sampled.

In the largest retrospective study included in this meta-analysis and testing the accuracy of PI-RADS v2 on 3.0 Tesla mpMRI in detecting csPCa with whole-mount sections as the reference standard, 2 readers with 9 and 4 years of experience in prostate mpMRI independently analyzed images of 425 patients.[16] Sensitivity at threshold score ≥ 4 was 0.77 for both readers, and inter-reader agreement was excellent in the per-patient analysis. Unfortunately, no per-lesion analysis was performed.

The only prospective study evaluating the performance of PI-RADS v2 on 3.0 Tesla mpMRI against whole-mount sections as the standard of reference used the index lesion on pathology as the main outcome.[17] In this study including 34 patients, the performance of 5 readers with largely different experiences was assessed. When considering the 2 most experienced readers, sensitivity for index lesion detection was 0.93 at PI-RADS v2 threshold score ≥ 3, and inter-reader agreement was good. When considering all lesions, sensitivity for the same readers dropped to 0.63. However, some controversy does exist in considering the index lesion as the biological driver of PCa progression and lethality.[18] We, thus, chose csPCa as a more appropriate outcome measure.

A matter of current debate is which PI-RADS v2 threshold should be used to trigger prostate biopsy. In our study, sensitivity of PI-RADS v2 using threshold score ≥ 3 vs ≥ 4 was comparable between readers and inter-reader agreement was similarly substantial, however FDR nearly halved for threshold score 4. As supported by data in Table 2, this suggests that PI-RADS v2 score 3 assignments do reflect ambiguous imaging findings, which ultimately translate in a too high proportion of false positive reads to be proposed as an effective threshold to prompt biopsy.

Of the missed cancers at mpMRI, most were nonclinically significant cancers, which is a desirable goal to limit the overdetection phenomenon of PCa screening and its negative consequences. A small proportion of them were, in fact, small-sized clinically significant foci, which would mean that, if confirmed by other studies, dimensional criteria to segregate certain PI-RADS categories should be critically reconsidered in future revision of the reporting system.

Our study is not devoid of limitations. First, all patients underwent RP, which may introduce a detection bias among readers because of the higher likelihood of finding intermediate- and high-risk PCa in this population of patients. In fact, an mpMRI diagnosis of csPCa may have been correct simply by chance, thereby artificially inflating detection rate and thus, sensitivity. Pretreatment assessment with mpMRI is, however, one of the most common clinical scenarios where PIRADSv2 is used. Second, the study was not enough powered to allow a reliable analysis of diagnostic performance of PI-RADS v2 to detect csPCa depending on the threshold score (3/4) or on tumor zonal location (peripheral zone vs transition zone). Third, we did not assess diagnostic accuracy measures on a wider range of reading experience, which might be of interest in validating PI-RADS v2 as a tool for less experienced or even naïve, readers. However, previous studies focusing on this topic did not find significant differences in accuracy between experienced and less experienced readers.[17,19]

CONCLUSIONS

In our prospective study with definitive pathology after RP as standard of reference, PI-RADS v2 showed good sensitivity in detecting csPCa on 3.0 Tesla mpMRI with substantial agreement between 2 experienced readers. Using PI-RADS v2 ≥4 as threshold for test positivity resulted in lower FDR compared with PI-RADS v2 ≥3,

which might have implications in setting a trigger for MRI-informed prostate biopsy.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.urology.2018.07.067.

References

1. Fütterer JJ, Briganti A, De Visschere P, et al. Can clinically significant prostate cancer be detected with multiparametric magnetic resonance imaging? A systematic review of the literature. *Eur Urol*. 2015;68:1045–1053.
2. Rosenkrantz AB, Ginocchio LA, Cornfeld D, et al. Interobserver reproducibility of the PI-RADS version 2 lexicon: a multicenter study of six experienced prostate radiologists. *Radiology*. 2016;280:793–804.
3. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol*. 2012;22:746–757.
4. Hamoen EHJ, de Rooij M, Witjes JA, Barentsz JO, Rovers MM. Use of the prostate imaging reporting and data system (PI-RADS) for prostate cancer detection with multiparametric magnetic resonance imaging: a diagnostic meta-analysis. *Eur Urol*. 2015;67:1112–1121.
5. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate imaging - reporting and data system: 2015, version 2. *Eur Urol*. 2016;69:16–40.
6. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic performance of prostate imaging reporting and data system version 2 for detection of prostate cancer: a systematic review and diagnostic meta-analysis. *Eur Urol*. 2017;72:177–188.
7. Epstein JI, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol*. 2016;40:244–252.
8. van der Kwast TH, Amin MB, Billis A, et al. International Society of Urological Pathology (ISUP) consensus conference on handling and staging of radical prostatectomy specimens. Working group 2: T2 substaging and prostate cancer volume. *Mod Pathol*. 2011;24:16–25.
9. Ploussard G, Epstein JI, Montironi R, et al. The contemporary concept of significant versus insignificant prostate cancer. *Eur Urol*. 2011;60:291–303.
10. Bains LJ, Studer UE, Froehlich JM, et al. Diffusion-weighted magnetic resonance imaging detects significant prostate cancer with high probability. *J Urol*. 2014;192:737–742.
11. Radtke JP, Schwab C, Wolf MB, et al. Multiparametric magnetic resonance imaging (MRI) and MRI-transrectal ultrasound fusion biopsy for index tumor detection: correlation with radical prostatectomy specimen. *Eur Urol*. 2016;70:846–853.
12. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57:289–300.
13. Fleiss JL, Levin B, Paik MC. Sample size for a one-sample study. In: Fleiss JL, Levin B, Paik MC, eds. *Statistical methods for rates and proportions*; 3rd Edition, Wiley, New York, 2003. pp.29–36.
14. Fleiss JL, Levin B, Paik MC. The measurement of interrater agreement. In: Fleiss JL, Levin B, Paik MC, eds. *Statistical methods for rates and proportions*; 3rd Edition, Wiley, New York, 2003. pp.598–626.
15. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. 2015;277:826–832.
16. Park SY, Jung DC, Oh YT, et al. Prostate cancer: PI-RADS version 2 helps preoperatively predict clinically significant cancers. *Radiology*. 2016;280:108–116.
17. Greer MD, Brown AM, Shih JH, et al. Accuracy and agreement of PIRADSv2 for prostate cancer mpMRI: a multireader study. *J Magn Reson Imaging*. 2017;45:579–585.
18. Giannarini G, Gandaglia G, Montorsi F, Briganti A. Will focal therapy remain only an attractive illusion for the primary treatment of prostate cancer? *J Clin Oncol*. 2014;32:1299–1301.
19. Sonn GA, Fan RE, Ghanouni P, et al. Prostate magnetic resonance imaging interpretation varies substantially across radiologists. *Eur Urol Focus*. 2017;pii: S2405-4569:30266–30303. http://dx.doi.org/10.1016/j.euf.2017.11.010. 2018 [Epub ahead of print].

EDITORIAL COMMENT

In this single-center study, Giannarini et al. prospectively evaluated the sensitivity values and false detection rates (FDRs) achieved by 2 experienced readers using the Prostate Imaging Reporting and Data System, version 2 (PI-RADS v2), for the detection of clinically significant cases of prostate cancer (PCa). The study included 48 patients who underwent 3 Tesla multiparametric magnetic resonance imaging (MRI) without endorectal coil before undergoing radical prostatectomy (RP). Whole-mount RP specimens were used as the reference standard, and clinically significant PCa was defined as tumors with any of the following characteristics: dimension ≥1 cm; International Society of Urological Pathology Grade Group ≥2; and pathologic stage ≥T3. As a secondary outcome, the authors also assessed the inter-reader agreement.

Sensitivity values were higher for patient-level analysis (73%-85%) than for lesion-level analysis (61%-75%). FDRs were significantly lower for both readers when PI-RADS v2 scores ≥4 instead of ≥3 were considered positive (0.06-0.10 vs 0.13-0.17, respectively), at the expense of a "slightly lower" sensitivity. The inter-reader agreement was substantial for PI-RADS v2 scores ≥3 and ≥4 (k = 0.72 and 0.65, respectively).

The results from this study add to the already large body of research and clinical experience that supports the value of MRI for the detection of clinically significant cases of PCa. Certain aspects of this study deserve to be highlighted. First, the study was performed in the context of a quality improvement program that involved routine MRI of the prostate before RP. The authors should be lauded for this initiative, as such practices are of paramount importance to validate PI-RADS v2 and to build local expertise. Second, the study used RP specimens as a reference standard, which is considered the "ground truth." Although the inclusion of only patients who underwent RP may have introduced some bias to the study, the use of lesion-level analysis should mitigate this limitation. Lastly, in terms of trade-offs between sensitivity and FDR, the results of this study suggest that a PI-RADS v2 score ≥4 could be adopted as a threshold to trigger biopsy in order to decrease FDRs without significantly compromising sensitivity for PCa detection. In a study conducted by our group, we found a similar trend for lesions located in the transition zone but not for lesions located in the peripheral zone.[1] However, it is worth noting that in the study by Giannarini et al., 10.8%-21.7% of the false-negative lesions using the threshold of PI-RADS v2 score ≥4 were not intracapsular. Based on previous studies showing that the probability of PCa detection incrementally increases with increasing PI-RADS v2 scores,[2] it is unlikely that a single threshold will be universally adopted to dictate the need for a biopsy. Instead, PI-RADS v2 scores will need to be combined with clinical variables such as

serum prostate-specific antigen levels, PSA density, and previous biopsy histology to determine the most appropriate diagnostic pathway, as indicated in a recent publication by the PI-RADS steering committee.[3]

Andrei S. Purysko, MD, Section of Abdominal Imaging and Nuclear Radiology Department, Imaging Institute, Cleveland Clinic, Cleveland, OH

References

1. Purysko AS, Bittencourt LK, Bullen JA, Mostardeiro TR, Herts BR, Klein EA. Accuracy and interobserver agreement for prostate imaging reporting and data system, version 2, for the characterization of lesions identified on multiparametric MRI of the prostate. *AJR*. 2017;209:339–349.
2. Greer MD, Shih JH, Lay N, et al. Validation of the dominant sequence paradigm and role of dynamic contrast-enhanced imaging in PI-RADS version 2. *Radiology*. 2017;285:859–869.
3. Padhani AR, Weinreb J, Rosenkrantz AB, Villeirs G, Turkbey B, Barentsz J. Prostate imaging-reporting and data system steering committee: PI-RADS v2 status update and future directions. *Eur Urol* 2018; [Epub ahead of print].

https://doi.org/10.1016/j.urology.2018.07.068
UROLOGY 123: 196–197, 2019.

AUTHOR REPLY

We thank the authors for their positive notes on our study. We agree with them that prostate lesions in the peripheral zone should probably be characterized on multiparametric MRI and clinically managed subsequently in a different manner compared to those in the transition zone. It is a fact that imaging accuracy for the latter ones is still far from optimal. We move a step further by saying that prostate lesions should probably be differently characterized on multiparametric MRI depending on their location (eg, apex vs base) even when lying in the same (peripheral) zone. The possibility for a region-dependent Prostate Imaging Reporting and Data System with a different threshold for biopsy has already been alluded to by prominent experts in the field.[1]

We also agree with the authors of this comment that the sole characterization of prostate lesions on multiparametric MRI should not trigger a biopsy, but this should come after a thorough multifactorial risk assessment and tailored counseling on an individual basis. This holds true especially considering the potential implications of the recent randomized PRostate Evaluation for Clinically Important Disease: Sampling Using Image-guidance Or Not? (PRECISION) trial, where upfront multiparametric MRI followed by MRI-targeted biopsy in biopsy-naïve men allowed for a significantly higher rate of detection of clinically important prostate cancer compared to standard biopsy with no MRI.[2] The results of this study might lead to an epochal shift in the management of men referred to early detection of prostate cancer, where MRI might be liberally ordered even by nonurologists to any-risk individual. Clearly, not all men would benefit from an upfront MRI. Truong et al. developed and prospectively validated a calculator to predict the pretest probability of detecting high-risk prostate lesions on multiparametric MRI using age, prostate-specific antigen level, and prostate volume as input variables.[3] Tools like this should help heath care providers and patients make an informed decision on whether to undergo an upfront MRI.

Gianluca Giannarini, MD, Rossano Girometti, MD, Alessandro Crestani, MD, Vincenzo Ficarra, MD, the Urology Unit, Academic Medical Centre Hospital "Santa Maria della Misericordia", Udine, Italy; the Department of Medicine, Institute of Radiology, University of Udine, Udine, Italy; the Department of Human and Pediatric Pathology "Gaetano Barresi", Urologic Section, University of Messina, Messina, Italy

References

1. Schouten MG, van der Leest M, Pokorny M, et al. Why and where do we miss significant prostate cancer with multi-parametric magnetic resonance imaging followed by magnetic resonance-guided and transrectal ultrasound-guided biopsy in biopsy-naïve men? *Eur Urol*. 2017;71:896–903.
2. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *N Engl J Med*. 2018;378:1767–1777.
3. Truong M, Baack Kukreja JE, Rais-Bahrami S, et al. Multi-institutional clinical tool for predicting high-risk lesions on 3 tesla multiparametric prostate magnetic resonance imaging. *Eur Urol Oncol* 2018 [Epub ahead of print].

https://doi.org/10.1016/j.urology.2018.07.069
UROLOGY 123: 197, 2019.