The Accuracy of Prostate Magnetic Resonance Imaging Interpretation: Impact of the Individual Radiologist and Clinical Factors

Nicholas A. Pickersgill, Joel M. Vetter, Neel S. Raval, Gerald L. Andriole, Anup S. Shetty, Joseph E. Ippolito, and Eric H. Kim

OBJECTIVE To compare test performance of multiparametric magnetic resonance imaging (mpMRI) for detection of prostate cancer between individual radiologists using the Prostate Imaging Reporting and Data System (PI-RADS) and to identify clinical factors that may predict test performance.

MATERIALS AND METHODS We examined our database of consecutive men who received prostate mpMRI prior to biopsy between September 2014 and December 2016 (n = 459). Test performance (eg, sensitivity, specificity, positive predictive value [PPV] and negative predictive value) were defined with PI-RADS classification 4 or 5 considered test positive and Gleason score ≥7 on biopsy from any targeted core considered outcome positive. Multivariate logistic regression was performed to identify clinical variables that affect test performance.

RESULTS No significant differences in test performance were found among individual radiologists. Prior biopsy (odds ratio [OR] 0.10, $P = .01$), radiologist experience >500 prostate mpMRI (OR 0.18, $P = .04$), transition zone location (OR 0.10, $P = .04$), and posterior location (OR 0.04, $P = .03$) were predictors of diminished sensitivity. Location of the mpMRI lesion in the TZ was a predictor of improved specificity (OR 2.53, $P = .04$). Increasing age (OR 1.07, $P < .01$) and prostate-specific antigen (OR 1.10, $P < .01$) predicted increased PPV, while prior biopsy predicted decreased PPV (OR 0.50, $P < .01$).

CONCLUSION Although variation exists in test performance among individual radiologists using PI-RADS, significant differences were not observed. Additional prostate mpMRI experience was not beneficial in improving accuracy of interpretation. Nonmodifiable patient variables—including prostate lesion location, prior biopsy history, prostate-specific antigen, and age—are predictive of prostate mpMRI test performance. UROLOGY 127: 68–73, 2019. © 2019 Elsevier Inc.

Prostate multiparametric magnetic resonance imaging (mpMRI) has become an important tool in the management of prostate cancer (PCa). The use of mpMRI to identify patients who require a biopsy as well as to target specific suspicious mpMRI lesions during biopsy has improved the detection of clinically significant PCa.[1-3] In order to standardize the interpretation of prostate mpMRI, the Prostate Imaging-Reporting and Data System (PI-RADS) was created, and recently updated to its second version.[4]

Although prostate mpMRI targeted biopsy (MRITB) is a promising and continuously improving technology, it has important limitations.[5-7] A significant driver of prostate mpMRI test performance is the individual radiologist interpreting the study. Despite the standardization provided by PI-RADS, significant variability in radiologists' interpretations of suspicious prostate lesions has been reported.[8-11] Given this variability in interpretation, one would also expect to observe differences in measures of test performance (eg, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)) between individual radiologists. The extent to which these differences exist, however, is unknown.

Most published studies of prostate mpMRI have utilized single, "expert" readers in a research setting. Moreover, few studies of interobserver variability have compared interpretations between reviewers with direct correlation to the pathologic findings. As prostate mpMRI continues to be more widely adopted, clinical practice settings will

Financial Disclosures: None.
Conflicts of Interest: None.
Funding/Support: We did not receive funding specifically for this study.

From the Division of Urology, Washington University School of Medicine, St. Louis, MO; and the Department of Radiology, Washington University School of Medicine, St. Louis, MO

Address correspondence to: Eric H. Kim, M.D., Division of Urologic Surgery, Washington University School of Medicine, 4960 Children's Place, Campus Box 8242, St. Louis, MO 63110. E-mail: ehkim@wustl.edu

Submitted: October 9, 2018, accepted (with revisions): January 16, 2019

https://doi.org/10.1016/j.urology.2019.01.035
0090-4295

less likely have a single, high-volume radiologist as a dedicated reader for all images, and test performance in relation to pathology will be more clinically meaningful than interobserver variability in isolation. Therefore, identifying the variability of test performance among radiologists of varying levels of experience with interpreting prostate mpMRI in a clinical setting becomes important. At our institution, prostate mpMRI are read as part of the clinical workflow of the abdominal imaging section, resulting in single reads for each study by 1 attending radiologist. Using MRITB outcomes, we compare the test performance (sensitivity, specificity, PPV, and NPV) of prostate mpMRI interpretation between 9 radiologists using PI-RADS at our institution. Additionally, we aim to identify clinical factors that may predict or influence prostate mpMRI test performance.

MATERIALS AND METHODS

With Institutional Review Board approval, we retrospectively examined our database of consecutive men who underwent prostate mpMRI following the implementation of PI-RADS by our institution in September 2014. Between September 2014 and December 2016, we identified 459 patients who underwent prostate mpMRI prior to biopsy. A total of 9 attending radiologists interpreted the 459 studies, with each radiologist having read a minimum of 11 studies. Patients with PI-RADS classification 1 or 2 lesions underwent standard template transrectal ultrasound (TRUS) guided biopsies, while patients with PI-RADS classification 3 to 5 lesions underwent both standard template TRUS-guided biopsy as well as software-fusion MRITB. The UroNav (Invivo Corporation, Gainesville, FL) platform was used to perform software-fusion MRITB. In all patients, 4-5 biopsy cores were obtained from each MRI target lesion and 1-2 biopsy cores were obtained from each sextet region. All biopsies were performed by a single attending urologist (GLA). Tissue biopsy cores were evaluated histopathologically as part of the clinical workflow of the anatomic pathology section. The highest Gleason score on biopsy (for both targeted and systematic cores) was used for analysis in order to eliminate technical error as a source of confounding. Gleason score ≥7 was considered clinically significant PCa.

MRI Technique

Specific description of our institutional prostate mpMRI technique has been published previously.[7,12] Briefly, all patients underwent 3-Tesla mpMRI using a pelvic phased-array coil on Siemens Trio and Skyra platforms (Siemens Healthcare, Erlangen, Germany). Patients received intravenous gadolinium contrast at 0.1 mmol/kg infused at a rate of 2 mL/s. High-resolution turbo spin echo T2-weighted images consisted of 3-mm slice thickness, small field of view imaging (160 mm) with a matrix of 512 × 512 for axial imaging and 640 × 640 for coronal imaging. The diffusion-weighted imaging sequence initially consisted of 3 b-values: 50, 500, and 1000 with a matrix of 128 × 128 and 3-mm slices, and more recently b-values of 50 and 800 with a calculated b-value of 1400. Dynamic contrast enhancement sequences consisted of gradient recall echo with a temporal resolution of 6-8 seconds, imaged initially over 6 minutes, later shortened to 2 minutes, with a matrix of 256 × 256 and 3-mm slice thickness.

MRI Interpretation

Prostate mpMRI were interpreted as part of an academic radiology practice by a group of 9 fellowship-trained radiologists with 2-11 years of experience in interpretation of prostate mpMRI, working with residents and fellows. Prostate mpMRI assignments were determined by the clinical workflow within the abdominal imaging section. Radiologist experience was measured by number of years in practice following training and the number of prostate mpMRIs interpreted both during and after training. PI-RADS version 1 was implemented at our institution in September 2014, followed by PI-RADS version 2 in February 2015. Although there are differences between the earlier and current scoring systems, recent studies have shown excellent inter-reader agreement for both scoring systems.[13]

Statistical Analysis

Sensitivity, specificity, PPV, and NPV, herein referred to collectively as "test performance," were calculated using 2 × 2 contingency tables individually for each of the nine attending radiologists. PI-RADS classification 1-3 versus 4-5 was considered test negative versus test positive, respectively, consistent with PI-RADS assessment category specifications.[4] Separate analyses of test performance were conducted using pathology from either the fusion (targeted) or systematic (random) biopsy cores as reference standards. Biopsy Gleason score <7 versus ≥7 was considered condition negative versus condition positive, respectively. Sensitivity analyses were performed with various combinations of threshold PI-RADS classifications and Gleason scores, of which the above-mentioned values resulted in the strongest test performance. Multivariate logistic regression analysis was performed to determine the effect of clinical variables on the measures of test performance. All analyses were performed with R version 3.2.2. *P* values below .05 were considered significant.

RESULTS

Patient characteristics and mpMRI results are summarized in Table 1. The mean age of our cohort was 64.8 ± 7.4 years, mean prostate-specific antigen (PSA) was 8.7 ± 10.5 ng/mL, and digital rectal examination was positive in 13.3% (61/459) of patients. Of the 459 patients included, 18.1% (83) were on active surveillance for previously diagnosed PCa, 40.1% (184) had a prior negative biopsy, and 41.8% (192) were biopsy-naïve. In all patients who had undergone prior biopsy, prostate MRI was performed at least 12 weeks following the biopsy. The mean number of biopsy cores per patient was 16. MRI lesions (of any PI-RADS score) were visualized in 86.5% (397/459) of patients. Within this subset of patients, mean MRI lesion volume was 1.5 ± 2.4 mL, 22.4% (89/397) had anterior lesions, and 80.1% (318/397) had lesions located exclusively in the peripheral zone. Clinically significant PCa (Gleason ≥7) was detected histologically in 15.9% of negative mpMRI, 25.0% of PI-RADS 1, 13.0% of PI-RADS 2, 11.9% of PI-RADS 3, 34.5% of PI-RADS 4, and 62.9% of PI-RADS 5 lesions.

The mean number of prostate mpMRIs interpreted by each attending radiologist during and after training was 535 ± 300, and the mean length of post-training experience for each attending radiologist was 11.3 ± 9.3 years.

Using pathology from fusion (targeted) biopsy cores as the reference standard, sensitivity, specificity, PPV, and NPV were 83.6%, 50.5%, 41.0%, and 88.2%, respectively. Using pathology from systematic (random) biopsy cores, sensitivity, specificity, PPV, and NPV were 82.9%, 47.9%, 46.2%, and 86.0%,

Table 1. Patient characteristics and mpMRI results

Characteristic	Value
Age (y), mean (SD)	64.8 (7.4)
PSA (ng/mL), mean (SD)	8.7 (10.5)
Positive DRE, n (%)	61/459 (13.3)
Race, n (%)	
Caucasian or other	420/459 (91.5)
African-American	39/459 (8.5)
Family history of PCa, n (%)	
None or unknown	324/459 (70.6)
Positive	135/459 (29.4)
Prostate volume (mL), mean (SD)	57.0 (29.1)
Biopsy history, n (%)	
Prior positive (active surveillance)	83/459 (18.1)
Prior negative	184/459 (40.1)
Naïve	192/459 (41.8)
MRI lesions, n (%)	
No lesion/PI-RADS 1-2	91/459 (19.8)
PI-RADS 3	95/459 (20.7)
PI-RADS 4-5	273/459 (59.5)
Anterior	89/397 (22.4)
Peripheral zone	318/397 (80.1)
MRI lesion volume (mL), mean (SD)	1.5 (2.4)
Biopsy cores per patient, mean	16

DRE, digital rectal exam; MRI, magnetic resonance imaging; PCa, prostate cancer; PI-RADS, Prostate Imaging Reporting and Data System; PSA, prostate-specific antigen.

respectively. Overall accuracy, a composite value of PPV and NPV, was 60.1% and 62.3% when using pathology from fusion and systematic biopsy cores, respectively. Between individual radiologists, sensitivity ranged from 71% to 100%, specificity ranged from 30% to 63%, PPV ranged from 36% to 56%, and NPV ranged from 78% to 100%. Measures of test performance did not differ significantly between individual radiologists, and were not correlated with measures of radiologist experience, such as total number of prostate mpMRI interpreted or years of experience.

On multivariate analysis, prior biopsy history (odds ratio [OR] 0.10, $P = .01$), radiologist experience >500 prostate mpMRI (OR 0.18, $P = .036$), and transition versus peripheral zone lesion location (OR 0.10, $P = .040$) were all independently predictive of diminished sensitivity, while anterior lesion location was independently predictive of increased sensitivity (OR 25.0, $P = .03$). Transition versus peripheral zone lesion location was independently predictive of increased specificity (OR 2.53, $P = .04$). Increasing age (OR 1.07, $P < .01$), PSA (OR 1.10, $P \leq .01$), and MRI lesion volume (OR 1.27, $P = .02$) were all independently associated with increased PPV, while prior biopsy was independently associated with reduced PPV (OR 0.37, $P < .01$). Radiologist experience >500 prostate mpMRI was independently associated with diminished NPV (OR 0.18, $P = .03$). Results of multivariate logistic regression analyses are shown in Table 2.

Table 2. Multivariate logistic regression analysis of clinical factors predictive of measures of test performance

Sensitivity		
Variable	Odds Ratio (95% CI)	*P* value
Age (continuous)	1.13 (1.00, 1.28)	.057
PSA (continuous)	1.28 (0.97, 1.70)	.082
Prior biopsy	0.10 (0.02, 0.58)	.010
Radiologist experience >500 MRIs	0.18 (0.04, 0.89)	.036
Transition zone	0.10 (0.01, 0.90)	.040
Anterior location	25 (1.45, 434.8)	.027
Lesion MRI volume	1.11 (0.58, 2.14)	.749
Specificity		
Variable	Odds Ratio (95% CI)	*P* value
Age (continuous)	0.99 (0.95, 1.03)	.613
PSA (continuous)	0.92 (0.84, 1.01)	.067
Prior biopsy	0.88 (0.48, 1.62)	.683
Radiologist experience >500 MRIs	0.95 (0.52, 1.73)	.872
Transition zone	2.53 (1.04, 6.18)	.042
Anterior location	0.41 (0.16, 1.06)	.068
Lesion MRI volume	0.85 (0.66, 1.11)	.230
PPV		
Variable	Odds Ratio (95% CI)	*P* value
Age (continuous)	1.07 (1.02, 1.11)	.005
PSA (continuous)	1.10 (1.03, 1.17)	.003
Prior biopsy	0.37 (0.20, 0.69)	.002
Radiologist experience >500 MRIs	1.20 (0.65, 2.22)	.567
Transition zone	0.77 (0.27, 2.17)	.618
Anterior location	1.70 (0.72, 4.00)	.222
Lesion MRI volume	1.27 (1.05, 1.55)	.015
NPV		
Variable	Odds Ratio (95% CI)	*P* value
Age (continuous)	0.99 (0.90, 1.10)	.918
PSA (continuous)	1.06 (0.87, 1.29)	.564
Prior biopsy	0.29 (0.06, 1.32)	.109
Radiologist experience >500 MRIs	0.18 (0.04, 0.81)	.026
Transition zone	0.40 (0.09, 1.81)	.232
Anterior location	0.34 (0.08, 4.37)	.600
Lesion MRI volume	0.76 (0.39, 1.48)	.419

CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

DISCUSSION

In our experience with PI-RADS for prostate mpMRI interpretation in a clinical practice setting, we find variation in measures of test performance among individual radiologists: sensitivity ranged from 71% to 100%, specificity ranged from 30% to 63%, PPV ranged from 36% to 56%, and NPV ranged from 78% to 100%. Interestingly, despite this range of test performance, no significant variability was observed between individual radiologists. Importantly, our overall institutional prostate mpMRI interpretations demonstrate stratification for biopsy-proven, clinically significant PCa that is comparable to "expert" readers in a research setting.[2]

Variation in diagnostic accuracy between radiologists has been investigated for imaging modalities in other diseases, most notably for breast cancer. Significantly diminished accuracy for mammogram interpretation has been reported for radiologists who read less than 500 mammograms per year, as well as those who more recently completed training.[14,15] Though we also find variation in test performance among radiologists at our institution, no

significant associations were noted with radiologist experience, neither in number of prostate mpMRI previously interpreted nor in years of practice. Interestingly, we found that prior interpretation of greater than 500 prostate mpMRI was predictive of reduced sensitivity and NPV with no corresponding increase in specificity or accuracy. As the use of prostate mpMRI is relatively novel, particularly outside of a research setting, increasing experience may not provide a benefit in the accurate interpretation of these studies. Furthermore, the standardization provided by PI-RADS may attenuate the advantages of experience, as Schieda et al found that more-experienced radiologists performed significantly better when using nonstandardized reporting in predicting extraprostatic extension at prostatectomy but performed similarly to less-experienced radiologists when using PI-RADS classification.[16]

In our multivariate analysis, we found that prior biopsy history was predictive of reductions in both sensitivity and PPV of prostate mpMRI. One potential explanation for these observed reductions in test performance is the known postbiopsy artifact caused by intraprostatic hemorrhage, which has been shown to cause inaccurate prostate mpMRI interpretation.[4,6,17] Although all patients in our cohort underwent prostate mpMRI at least 12 weeks following their previous biopsy (beyond the recommended 6 weeks),[4] artifact due to hemorrhage can require up to 4 months to completely resolve, thereby potentially impairing image interpretation.[17]

In our institutional experience prior to the implementation of PI-RADS, we found that mpMRI was not predictive of clinically significant PCa in those with previous negative biopsies, while mpMRI was predictive for those who were biopsy naïve.[18] As the majority of those with prior biopsy in our cohort had a prior negative biopsy (69%), the reductions in test performance when using PI-RADS may also be for similar reasons to our experience. Additionally, given that our radiologists were not blinded to patient clinical data, it is possible that knowledge of prior biopsy status may have imposed clinical bias on study interpretations. For example, given the subtle differences between PI-RADS 3 and 4 classifications (and our study definition of PI-RADS 3 as test negative and PI-RADS 4 as test positive, consistent with the PI-RADS lexicon),[4] the known history of a previous negative biopsy could result in a downgrade from PI-RADS 4 to 3, thereby generating a false negative result and, consequently, decreasing both sensitivity and accuracy. Nonetheless, these results may support an important role of prostate mpMRI in the diagnostic workup of biopsy-naïve men for clinically significant PCa.

Our analysis found that transition zone lesion location was predictive of both decreased prostate mpMRI sensitivity and specificity. The limited ability of mpMRI to diagnose transition zone PCa, due to the heterogenous T2 signal intensity in the normal transition zone, has been well documented.[19] While T2-weighted imaging is considered the most accurate of the mpMRI sequences for the detection of transition zone PCa, they remain difficult to differentiate from fibromuscular benign prostatic hyperplasia as both conditions can demonstrate similarly low T2 signal intensity.[19-21] With this in mind, our study finding of transition zone lesions being predictive of diminished sensitivity and specificity is expected. Conversely, prostate mpMRI is known to play a crucial role in identifying and locating anterior tumors, which are relatively difficult to diagnose with conventional TRUS-guided biopsy.[22,23] Given this known benefit, the improved sensitivity associated with anterior lesion location is not surprising. Considering the high rate of anterior PCa detection in early prostate mpMRI series,[23] the radiologists may have been more careful in examining the anterior portions of the prostate during their mpMRI interpretation.

The increased likelihood of diagnosing clinically significant (Gleason ≥7) PCa with increasing prebiopsy PSA has been well established.[24,25] Notably, the PPV of PSA for PCa has been found to increase significantly when using 10.0 ng/mL rather than 4.0 ng/mL as a threshold for biopsy.[24,26] Although the measurement of PSA is a critical part of the PCa diagnostic pathway, much controversy exists surrounding its use as a screening tool.[27] However, the use of prostate mpMRI to triage biopsy-naïve men may avoid unnecessary TRUS-guided biopsy and the over-diagnosis of clinically insignificant PCa in men with elevated PSA.[28] The prevalence of PCa is also known to be strongly linked to increasing age, with the prevalence doubling approximately every 14 years of life.[29] Similarly, our study found that increasing age was predictive of increased PPV of prostate mpMRI. Of note, this result is likely confounded by the increased prevalence of PCa with increasing age.

The main limitation of our study is its retrospective design, as the analysis of clinical data generates biases that cannot be completely accounted for. However, as a descriptive study of prostate mpMRI in a clinical setting, the study results are likely more representative of radiology practices across the country, thus providing greater insights into factors affecting the test performance of prostate mpMRI interpretation when using PI-RADS. Second, the impact of radiologists remaining unblinded to clinical parameters cannot be quantified; however, this is likely more representative of how mpMRI will be interpreted outside of a research setting. Ultimately, a prospective assessment of "non-expert" radiologists, blinded to patient clinical data, with multiple radiologists reading the same mpMRI in an independent fashion with correlation to either biopsy or prostatectomy pathology is necessary. An additional limitation of our study is the implementation of PI-RADS version 2 during the timeframe of observation. However, as stated previously, recent studies have shown excellent inter-reader agreement and comparable diagnostic performance between the 2 versions, suggesting that this transition would have minimal impact on study results.[13]

CONCLUSION

Although variability in test performance exists among individual radiologists interpreting prostate mpMRI using PI-RADS in a clinical setting, no significant differences in sensitivity, specificity, PPV, or NPV were observed. Additional radiologist experience, measured by years in practice or number of prostate mpMRI interpreted, is not predictive of improved test performance. Instead, nonmodifiable patient variables—namely prostate lesion location and prior biopsy history—are predictive of prostate mpMRI test performance.

References

1. 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.
2. Rastinehad AR, Turkbey B, Salami SS, et al. Improving detection of clinically significant prostate cancer: magnetic resonance imaging/transrectal ultrasound fusion guided prostate biopsy. *J Urol*. 2014;191:1749–1754.
3. Tompson JE, Moses D, Shnier R, et al. Multiparametric magnetic resonance imaging guided diagnostic biopsy detects significant prostate cancer and could reduce unnecessary biopsies and over detection: a prospective study. *J Urol*. 2014;192:67–74.
4. Weinreb JC, Barentsz JO, Choyke PL, et al. Prostate imaging—reporting and data system: 2015, version 2. *Eur Urol*. 2016;69:16–40.
5. Nassiri N, Natarajan S, Margolis DJ, Marks LS. Targeted prostate biopsy: lessons learned midst the evolution of a disruptive technology. *Urology*. 2015;86:432–438.
6. Rosenkrantz AB, Taneja SS. Radiologist, be aware: ten pitfalls that confound the interpretation of multiparametric prostate MRI. *Am J Roentgenol*. 2014;202:109–120.
7. Wang RS, Kim EH, Vetter JM, et al. Determination of the role of negative magnetic resonance imaging of the prostate in clinical practice: is biopsy still necessary? *Urology*. 2017;102:190–197.
8. Riney JC, Sarwani NE, Siddique S, Raman JD. Prostate magnetic resonance imaging: the truth lies in the eye of the beholder. *Urol Oncol*. 2018;36. 159.e1-159.e5.
9. Rosenkrantz AB, Ginocchio LA, Cornfeld D, et al. Inter-observer reproducibility of the PI-RADS version 2 lexicon: a multi-center study of six experienced prostate radiologists. *Radiology*. 2016;280:793–804.
10. Rosenkrantz AB, Lim RP, Haghighi M, et al. Comparison of interreader reproducibility of the Prostate Imaging Reporting And Data System and Likert scales for evaluation of multiparametric prostate MRI. *Am J Roentgenol*. 2013;201:W612–W618.
11. Sonn GA, Fan RE, Ghanouni P, et al. Prostate magnetic resonance imaging interpretation varies substantially across radiologists. *Eur Urol Focus*. 2017;30266-30263. https://doi.org/10.1016/j.euf.2017/11/010.
12. Kim EH, Weaver JK, Shetty AS, et al. Magnetic resonance imaging provides added value to the prostate cancer prevention trial risk calculator for patients with estimated risk of high grade prostate cancer less than or equal to 10%. *Urology*. 2017;102:183–189.
13. Polanec S, Helbich TH, Bickel H, et al. Head-to-head comparison of PI-RADS v2 and PI-RADS v 1. *Eur J Radiol*. 2016;85:1125–1131.
14. Tan A, Freeman DH, Goodwin JS, Freeman JL. Variation in false-positive rates of mammography reading among 1067 radiologists: a population-based assessment. *Breast Cancer Res Treat*. 2006;100:309–318.
15. Theberge I, Chang SL, Vandal N, et al. Radiologist interpretive volume and breast cancer screening accuracy in a Canadian organized screening program. *J Natl Cancer Inst*. 2014;106(3). djt461.
16. Schieda N, Quon JS, Lim C, et al. Evaluation of the European Society of Urogenital Radiology (ESUR) PI-RADS scoring system for assessment of extra-prostatic extension in prostatic carcinoma. *Eur J Radiol*. 2015;84:1843–1848.
17. Ahmed HU, Kirham A, Arya M, et al. Is it time to consider a role for MRI before prostate biopsy? *Nat Rev Clin Oncol*. 2009;6:197–206.
18. Weaver JK, Kim EH, Vetter JM, et al. Presence of magnetic resonance imaging suspicious lesion predicts Gleason 7 or greater prostate cancer in biopsy-naïve patients. *Urology*. 2016;88:119–124.
19. Akin O, Sala E, Moskowitz CS, et al. Transition zone prostate cancers: features, detection, localization, and staging at endorectal MR imaging. *Radiology*. 2006;239:784–792.
20. Hoeks CM, Hambrock T, Yakar D, et al. Transition zone prostate cancer: detection and localization with 3-T multiparametric MR imaging. *Radiology*. 2013;266:207–217.
21. Li H, Sugimura K, Kaji Y, et al. Conventional MRI capabilities in the diagnosis of prostate cancer in the transition zone. *Am J Roentgenol*. 2006;186:729–742.
22. Lawrentschuk N, Haider MA, Daljeet N, et al. 'Prostatic evasive anterior tumours': the role of magnetic resonance imaging. *Brit J Urol*. 2010;105:1231–1236.
23. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. *JAMA*. 2015;313:390–397.
24. Catalona WJ, Richie JP, Ahmann FR, et al. Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630 men. *J Urol*. 2017;197:S200–S207.
25. Schröder FH, van der Cruijsen-Koeter I, de Koning HJ, Vis AN, Hoedemaeker RF, Kranse R. Prostate cancer detection at low prostate specific antigen. *J Urol*. 2000;163:806–812.
26. Coley CM, Barry MJ, Fleming C, Mulley AG. Early detection of prostate cancer. Part I: prior probability and effectiveness of tests. *Ann Intern Med*. 1997;126:394–406.
27. Shoag JE, Mittal S, Hu JC. Reevaluating PSA testing rates in the PLCO trial. *N Engl J Med*. 2016;374:1795–1796.
28. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet*. 2017;389:815–822.
29. Bell KJL, Del Mar C, Wright G, Dickinson J, Glasziou P. Prevalence of incidental prostate cancer: a systematic review of autopsy studies. *Int J Cancer*. 2015;137:1249–1757.

Editorial Comment

Recent prospective studies, including randomized clinical trials, have provided compelling evidence in favor of the use of multiparametric magnetic resonance imaging (MRI) for the detection of clinically significant prostate cancer. Nevertheless, questions remain regarding how these results compare to those obtained in routine clinical practice when MRI examinations are interpreted by readers with variable levels of experience.

This retrospective study was performed at a single academic center and included 459 patients who had undergone prostate MRI before a biopsy was performed. The MRI exams were interpreted by 1 of 9 radiologists with experience ranging from 2 to 11 years. They initially used the scoring system described in Prostate Imaging Reporting and Data System (PI-RADS) version 1 and later adopted PI-RADS version 2. The prostate biopsies were performed by a single urologist, and the biopsy results were used as the reference standard to determine the accuracy of prostate MRI for detection of clinically significant prostate cancer (Gleason score ≥7). The authors reported variations in test performance among radiologists, but no significant associations were noted with radiologist experience (with experience defined in terms of number of prostate MRI examinations previously interpreted or years of practice). The authors argued that the use

of a standardized reporting system may attenuate the advantages of experience. Although this is a reasonable explanation, it does not fully explain why a radiologist's experience of interpreting more than 500 examinations was predictive of reduced sensitivity and negative predictive value. Instead of using an arbitrary number of studies or years of practice to define the readers' experience, the authors might have considered assessing how the performance of the readers changed over time, especially after the adoption of PI-RADS version 2.

I commend the authors for reporting on their institutional experience. More studies similar to this one, conducted both in academic and nonacademic centers, are needed to establish benchmarks for prostate MRI performance. Such benchmarks can be used to develop quality improvement initiatives and to set appropriate expectations among health care payers, urologists, and patients.

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

https://doi.org/10.1016/j.urology.2019.01.036
UROLOGY 127: 72–73, 2019.