Current Problems in Diagnostic Radiology 48 (2019) 132–141

Current Problems in Diagnostic Radiology

journal homepage: www.cpdrjournal.com

Online Interactive Case-Based Instruction in Prostate Magnetic Resonance Imaging Interpretation Using Prostate Imaging and Reporting Data System Version 2: Effect for Novice Readers

Andrew B. Rosenkrantz, MD[a,*], Jovan Begovic, MD[a], Antonio Pires, MD[a], Eugene Won, MD[a], Samir S. Taneja, MD[b], James S. Babb, PhD[a]

[a] Department of Radiology, NYU Langone Health, New York, NY
[b] Department of Pathology, NYU Langone Health, New York, NY

Purpose: To assess the effect on reader performance of an interactive case-based online tutorial for prostate magnetic resonance imaging (MRI) interpretation using Prostate Imaging and Reporting Data System (PI-RADS).
Methods: An educational website was developed incorporating scrollable multiparametric prostate MRI examinations with annotated solutions based on PI-RADS version 2. Three second-year radiology residents evaluated a separate set of 60 prostate MRI examinations both before and after review of the online case material, identifying and scoring dominant lesions. These 60 examinations included 30 benign cases and 30 cases with a dominant lesion demonstrating Gleason score ≥3 + 4 tumor on fusion-targeted biopsy. The readers' pooled performance was compared between the 2 sessions using logistic regression and Wilcoxon signed rank tests.
Results: All readers completed the online material within four-hours. Review of the online material significantly improved sensitivity (from 57.8%-73.3%, $P = 0.003$) and negative predictive value (from 69.2%-78.2%, $P = 0.049$), but not specificity (from 70.0%-67.8%, $P = 0.692$) or positive predictive value (from 59.6%-64.7%, $P = 0.389$). Reader confidence (1-10 scale; 10 = maximal confidence) also improved significantly (from 5.6 ± 2.7 to 6.3 ± 2.6, $P = 0.026$). However, accuracy of assigned PI-RADS scores did not improve significantly (from 45.5%-53.3%, $P = 0.149$).
Conclusion: An online interactive case-based website in prostate MRI interpretation improved novice readers' sensitivity and negative predictive value for tumor detection, as well as readers' confidence. This online material may serve as a resource complementing existing traditional methods of instruction by providing a more flexible educational experience among a larger volume of learners. However, further more targeted educational initiatives regarding the proper application of PI-RADS remain warranted.

Introduction

Prostate magnetic resonance imaging (MRI) has improved substantially in diagnostic performance in recent years, related to advances in both hardware and software along with continued insights from radiologic-pathologic correlation studies.[1] These improvements have led to a growing spectrum of clinical applications of prostate MRI, including tumor detection and localization, biopsy guidance, risk stratification, staging, surgical planning, and guidance of focal therapy.[2] Recent studies support the cost-effectiveness of incorporation of prostate MRI into routine clinical practice.[3,4] Statements from the American Urological Association (AUA), National Comprehensive Cancer Network, and European Association of Urology, all recognize the potential clinical role of prostate MRI,[5-7] and the AUA now offers regular educational courses on MRI-targeted biopsy to the urological community.[8] Thus, the dissemination of prostate MRI into smaller and community practices is anticipated to increase in the coming years.

Despite these advances, prostate MRI has been challenged by substantial interobserver variation in interpretation and reporting.[9,10] Even when high-quality images are obtained, interpretation is difficult given the heterogeneous appearance of the prostate, potentially subtle tumors, and various anatomical findings, pathologic conditions, and artifacts that lead to false-positive and false-negative readings.[11,12] Such interpretation challenges pose an important hurdle to the successful widespread expansion of prostate MRI into community settings. To address this issue, an international expert panel developed the prostate imaging and reporting data system (PI-RADS),[13] for which version 2 (v2) was released online in January 2015.[14] PI-RADS v2 aimed to "promote global standardization and diminish variation in prostate MRI readings."[14]

PI-RADS v2 has been the subject of intense interest in the United States since its release,[15-20] with a growing number of centers embracing the scheme. Nonetheless, even PI-RADS v2 is subject to interobserver variability or potential misreads. Indeed,

Taneja is a consultant for Hitachi-Aloka and receives nonfinancial support from Biobot. Rosenkrantz receives royalties from Thieme Medical Publishers.

This work was supported by an AUR/RSNA Education Scholar Grant.

* Reprint requests: Andrew B. Rosenkrantz, MD, Department of Radiology, Center for Biomedical Imaging, NYU Langone Health, 660 First Ave, 3rd Floor, New York, NY 10016.
E-mail address: Andrew.Rosenkrantz@nyumc.org (A.B. Rosenkrantz).

https://doi.org/10.1067/j.cpradiol.2018.01.003
0363-0188/

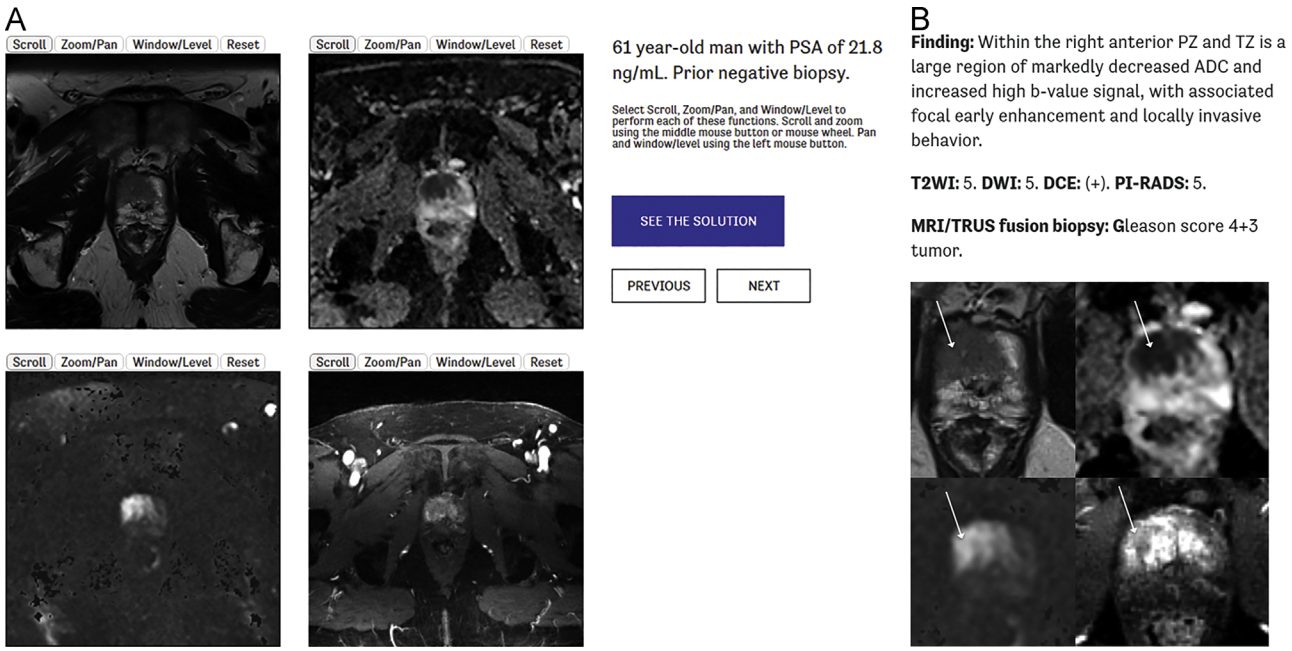

FIG 1. Representative screen shots of the interactive online educational platform. (A) Scrollable multiparametric data set. (B) Solution consisting of annotated images and explanation incorporating lesion assessment based on PI-RADS version 2.

an initial study of 6 experienced readers reported moderate reproducibility.[21] Although the system is comprehensive, further efforts are needed to educate radiologists regarding its nuances and optimal translation to clinical practice. Although high-volume centers may be rapidly gaining much direct experience in the system, the resulting insights must be disseminated through appropriately designed training opportunities to benefit the radiology community at large.[22,23] An educational resource meeting this goal, if successful, would be of immense value to radiologists both in practice and in training, throughout the United States and abroad.

Interactive case-based content with direct feedback may provide an effective means of learning PI-RADS v2. This approach provides individualized learning that engages participants and allows identification of areas needing greatest improvement. A prior study before the advent of PI-RADS demonstrated that interactive individually tailored dedicated training in prostate MRI led to significantly improved accuracy in comparison with standard didactic sessions alone.[24] An additional study demonstrated the role of a dedicated reader education program in substantially reducing interobserver variability.[25] Although these studies were effective in training participants at the authors' institutions, they were unable to directly benefit the broader radiology community. Web-based delivery of such content may achieve this end. Indeed, numerous studies have described the potential efficacy of web-based training for a variety of topics in radiology.[26-29]

The aim of this study was to assess the impact on reader performance of an interactive case-based online tutorial for prostate MRI interpretation using PI-RADS.

Methods

This retrospective study was health insurance portability and accountability act compliant and approved by the local institutional review board. A waiver of written informed consent was provided. The primary investigator (A.R.) is a fellowship trained abdominal radiologist with 9 years of experience in prostate MRI, currently serving on the PI-RADS Steering Committee.

Description of the Website

A publicly available website was developed featuring a library of 56 prostate MRI examinations.[30] For each examination, the website provides on a single screen 4 anonymized axial image sets as separate image stacks (Fig 1): T2-weighted imaging (T2WI), high *b* value diffusion-weighted imaging, apparent diffusion coefficient map, and an early time-point from dynamic contrast-enhanced imaging. The website incorporates a DICOM browser plug-in to allow the user to scroll through the image sets, as well as to window and level, pan, and zoom. Users may view the examinations as "unknowns," attempting to independently identify dominant prostate lesions and derive PI-RADS scores. Subsequently, the user may select to view an annotated solution depicting the dominant lesion, along with an explanation and associated PI-RADS scoring, before proceeding to the next case (Fig 1). The cases largely consist of histologically confirmed tumors of varying degrees of interpretive difficulty, though also include a smaller number of benign findings that mimic tumor, as well as normal examinations. In addition, the cases encompass the full spectrum of PI-RADS v2 scores in both the peripheral and transition zones.

Reader Evaluation

To assess the educational value of the website, a reader study was performed by 3 second-year radiology residents (A.A., B.B., C.C.) (J.B., A.P., and E.W.) who had not previously received formal instruction in prostate MRI interpretation during clinical rotations prior. The primary investigator provided the readers an introductory lecture covering the basics of prostate MRI interpretation. The readers also independently evaluated a small number of examinations not otherwise a part of the investigation to become familiar with the protocol and to provide an opportunity to ask questions regarding basic prostate anatomy and multiparametric sequence assessment.

An experimental design was constructed to assess for an improvement in diagnostic performance for a given set of prostate MRI examinations based on a review of the online case material:

(1) following the earlier described introduction, the three readers independently evaluated a set of 60 prostate MRI examinations that were distinct from the interactive website cases (hereafter, referred to as session 1). (2) Next, the readers were instructed to independently review the full set of website cases, viewing each case as an unknown prior to accessing the online annotated solution. (3) After completing the website review, the residents again evaluated the same initial set of 60 examinations (hereafter, referred to as session 2). For both sessions, the examinations were reviewed in an anonymized fashion (including removal of the clinical report and any image annotations) using the institutional PACS (iSite; Philips Healthcare). The readers were not provided solutions to the set of 60 test cases between the 2 sessions. In addition, during session 2, readers were not provided with their readings during session 1, which occurred at least 2 weeks earlier. Finally, when reviewing the online interactive case material, the readers were instructed not to access other portions of the website.

When evaluating the 60 cases during each session, the readers recorded whether or not a dominant lesion was identified. If so, the readers recorded the corresponding image and series number as well as the location of the lesion on a standardized diagram. An associated PI-RADS v2 assessment category was also recorded. In addition, for all cases, the readers recorded their confidence in their provided interpretation, whether positive or negative, based on a 1-10 scale (10 indicating maximal confidence).

The set of 60 cases was constructed to include 30 cases positive for tumor and 30 cases negative for tumor. Positive cases were drawn from the primary investigator's teaching files. All such cases met 3 criteria: (1) the MRI demonstrated a single dominant lesion for which the primary investigator assigned a PI-RADS v2 of 3 or higher; (2) MRI-ultrasound fusion-targeted biopsy of the lesion demonstrated tumor with a Gleason score of at least 3 + 4; and (3) systematic biopsy performed at the time of fusion biopsy did not demonstrate tumor at a location remote from the MRI lesion that had a Gleason score higher than 3 + 3. The 30 negative cases were selected at random from earlier investigations from our institution[31,32] and met the following criteria: (1) PI-RADS v2 score assigned by the lead investigator of 1 or 2 and (2) all subsequent systematic and targeted biopsies were benign. In addition, 19 of the patients were included in earlier unrelated examinations from our institution regarding outcomes of fusion-targeted biopsy.[33,34]

All MRI examinations were performed at 3T using a pelvic phased-array coil and included multi-planar T2WI, diffusion-weighted imaging with a calculated b value of 1500 s/mm^2, an apparent diffusion coefficient map, and dynamic contrast-enhanced at a temporal resolution of 2.3 seconds. The mean ± standard deviation (SD) patient age was 62.4 ± 6.9 years (median = 62 years; range: 51-79 years). The mean ± SD patient serum prostate-specific antigen was 7.6 ± 7.3 ng/mL (median = 5.9 ng/mL; range: 1.4-48.4 ng/mL). Indications for the MRI examinations were clinical suspicion for prostate cancer with no prior prostate biopsy (n = 31), prior negative prostate biopsy with persistent clinical suspicion for prostate cancer (n = 15), and prior biopsy positive for prostate cancer (n = 14).

Following completion of all reader evaluations, the primary investigator compared the readers' assessment with the histologic reference standard. For this purpose, reader scores were dichotomized, with assigned PI-RADS scores of 3-5 considered positive, and assigned PI-RADS scores of 1-2 considered negative. A spatial correspondence was required between the lesion called by the reader and the location of the lesion based on the reference standard to be deemed a match for purposes of defining a true-positive interpretation.

Survey

The readers completed a brief survey at 2 time-points during the study: (1) immediately after session 1 but before review of the online material and (2) immediately after review of the online material but before session 2. Thus, the only exposure to prostate MRI examinations between the 2 survey time-points consisted of the online case material itself. The prewebsite survey contained 3 statements regarding the readers' level of comfort and confidence in prostate MRI interpretation with which the readers rated their level of agreement on a scale of 1 (least agreement) to 5 (maximal agreement). The postwebsite survey contained these same 3 items, as well as an additional 9 items regarding satisfaction with the website and its perceived educational value with which the readers rated their level of agreement using the same scale. The postwebsite survey contained an additional three items regarding the appropriateness of the amount of content on the website and a free-response item regarding the amount of time needed to complete the online case material.

Statistics

Sessions 1 and 2 were compared in terms of the sensitivity, specificity, and overall accuracy of each reader using an exact McNemar test, and compared for the positive and negative predictive values (NPVs) of each reader, as well as for each component of diagnostic accuracy pooled over readers, using generalized estimating equations (GEE) based on a logistic regression model to predict accuracy as a function of session. These models treated data from the same patient as symmetrically correlated and data from different patients as independent. Similar comparisons of the 2 sessions were performed for sensitivity for select subsets of examinations based on lesion characteristics. The sessions were then compared to the accuracy of the readers' assigned PI-RADS scores, first among all examinations by considering the readers' scores as incorrect when failing to correctly detect the lesion, and then among those examinations in which a given reader correctly detected the tumor in both sessions. These comparisons in terms of accuracy of the PI-RADS score used an exact McNemar test for individual readers and used GEE when pooling across the three readers as previously described. Next, sessions 1 and 2 were compared in terms of individual readers' confidence scores using an exact paired-sample Wilcoxon signed rank test and for confidence scores pooled over readers using a stratified exact paired-sample Wilcoxon signed rank test with reader as the stratification factor. Then, inter-reader agreement for tumor detection and PI-RADS scores was assessed by comparing the percentage of times a pair of readers provided concordant interpretations using GEE as described previously. In addition, reader agreement in terms of tumor detection and PI-RADS scores were assessed using simple and linear weighted kappa coefficients, respectively, with the assessment of agreement for PI-RADS scores within each session being based on data for tumors correctly detected by a given pair of readers during both sessions. These assessments pooled data across the 3 distinct possible reader pairs. The level of agreement was interpreted as poor when kappa (κ) was less than zero, as slight when $0 \leq \kappa \leq 0.2$, as fair when $0.2 < \kappa \leq 0.4$, as moderate when $0.4 < \kappa \leq 0.6$ and as substantial when $\kappa > 0.6$. All statistical tests were conducted at the 2-sided 5% significance level using SAS 9.4 (SAS Institute, Cary, NC). Finally, the survey results were summarized descriptively without formal significance testing given the presence of only three readers.

TABLE 1
Comparison of measures of diagnostic performance between sessions 1 and 2

Measure	Reader	Session 1		Session 2		P^*
		%	*n*	%	*n*	
Sensitivity	1	60.0	18/30	83.3	25/30	**0.016**
	2	63.3	19/30	73.3	22/30	0.375
	3	50.0	15/30	63.3	19/30	0.289
	All	57.8	52/90	73.3	66/90	**0.003**
Specificity	1	66.7	20/30	70.0	21/30	1.000
	2	86.7	26/30	73.3	22/30	0.344
	3	56.7	17/30	60.0	18/30	1.000
	All	70.0	63/90	67.8	61/90	0.692
NPV	1	64.5	20/31	80.8	21/26	0.038
	2	74.3	26/35	81.5	22/27	0.279
	3	68.0	17/25	72.0	18/25	0.635
	All	69.2	63/91	78.2	61/78	**0.049**
PPV	1	65.5	19/29	73.5	25/34	0.217
	2	76.0	19/25	66.7	22/33	0.371
	3	42.9	15/35	54.3	19/35	0.190
	All	59.6	53/89	64.7	66/102	0.389
Accuracy	1	63.3	38/60	76.7	46/60	0.057
	2	75.0	45/60	73.3	44/60	1.000
	3	53.3	32/60	61.7	37/60	0.359
	All	63.9	115/180	70.6	127/180	0.096

* Highlighted in bold when statistically significant at $P < 0.05$.

Results

In the 30 examinations that were positive for tumor, the tumor's Gleason scores were as follows: 3 + 4 ($n = 14$), 4 + 3 ($n = 7$), 4 + 4 ($n = 1$), and 4 + 5 ($n = 8$). A total of 22 tumors were in the peripheral zone, and 8 were in the transition zone. The reference standard PI-RADS scores assigned to the lesions by the primary investigator were PI-RADS 3 ($n = 7$), PI-RADS 4 ($n = 12$), and PI-RDAS 5 ($n = 11$).

Table 1 compares measures of diagnostic performance between sessions 1 and 2 (before and after review of the online interactive cases, respectively). Pooled across the three readers, performance improved significantly from session 1 to session 2 in terms of sensitivity (from 57.8%-73.3%, respectively; $P = 0.003$) and NPV (from 69.2%-78.2% respectively; $P = 0.049$). There was no significant difference for accuracy (from 63.9%-70.6%, respectively; $P = 0.096$), specificity (from 70.0%-67.8%, respectively; $P = 0.692$), or positive predictive value (from 59.6%-64.7%, respectively; $P = 0.389$). Figures 2-4 demonstrate representative examples.

Table 2 compares the 2 sessions for sensitivity for selected subsets of tumors. Pooled across the three readers, sensitivity improved significantly from session 1 to session 2 for tumors with a Gleason score of 3 + 4 (from 48.9%-71.1%, respectively; $P = 0.006$) and for tumors located in the peripheral zone (from 60.6%-80.3%, respectively; $P < 0.001$). Improvements in sensitivity for tumors with a reference standard PI-RADS score of 3 (from 14.3%-42.9%, respectively; $P = 0.070$) and a reference standard PI-RADS score of 4 (from 61.1%-77.8%, respectively; $P = 0.07$) were not significant. There was also no significant difference in sensitivity for tumors with a PI-RADS score of 5, a Gleason score ≥4 + 3, or a location in the transition zone ($P \geq 0.289$).

Table 3 compares performance between sessions in providing a correct PI-RADS score. Pooled across readers, the accuracy of the assigned PI-RADS score in session 1 compared with session 2, for all tumors was 43.3% vs 53.3%, respectively ($P = 0.149$), for

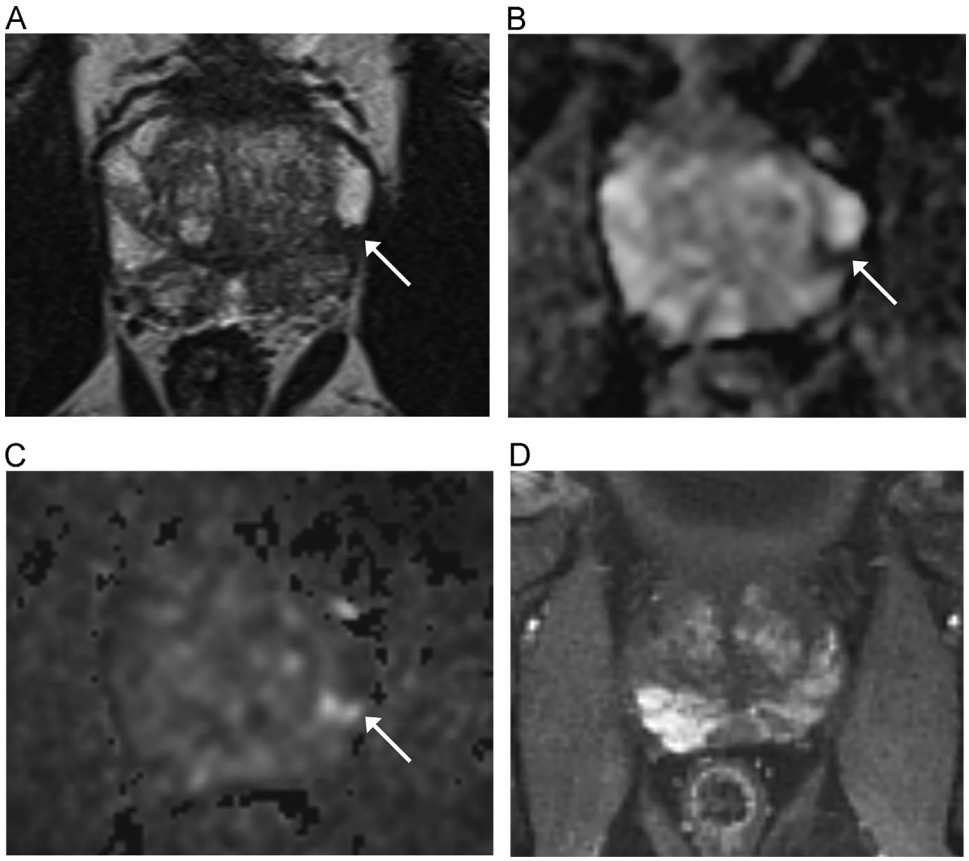

FIG 2. A 64 year-old man with Gleason score 3 + 4 tumor on MRI-ultrasound fusion-targeted biopsy of left peripheral zone lesion. The lesion was correctly detected by one reader in session 1, but by all 3 readers in session 2 following review of the online interactive case material.

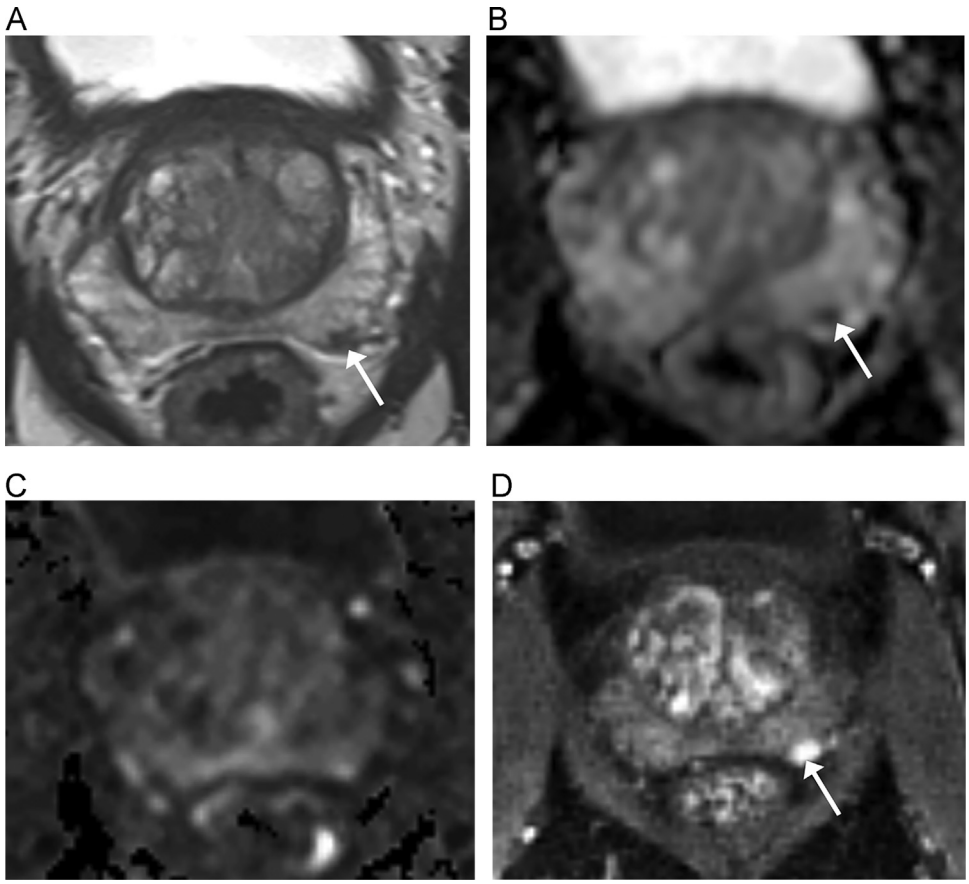

FIG 3. A 64 year-old man with Gleason score 4 + 3 tumor on MRI-ultrasound fusion-targeted biopsy of left peripheral zone lesion. The lesion was correctly detected by no readers in session 1, but by two of the three readers in session 2 following review of the online interactive case material.

tumors in the peripheral zone was 45.5% vs 59.1%, respectively (0.078), and for tumors in the transition zone was 37.5% in both sessions ($P = 1.0$). Comparisons of PI-RADS scores between sessions were not significant for subsets of examinations in which the reader correctly detected the tumor in both sessions ($P \geq 0.508$).

Table 4 compares reader confidence between the 2 sessions. Pooled across readers, confidence increased significantly from session 1 to session 2 for all examinations (from 5.6 ± 2.7-6.3 ± 2.6, respectively; $P = 0.026$), examinations positive for tumor (from 4.9 ± 2.5-5.7 ± 2.5, respectively; $P = 0.009$), examinations positive for Gleason score 3 + 4 tumor (from 6.0 ± 2.6-6.8 ± 2.5, respectively; $P = 0.026$), examinations positive for Gleason score ≥4 + 3 tumor (from 6.6 ± 2.7-7.0 ± 2.4, respectively; $P = 0.026$), examinations negative for tumor (from 4.9 ± 2.5-5.7 ± 2.5, respectively; $P = 0.004$), examinations positive for peripheral zone tumor (from 6.4 ± 2.8-7.0 ± 2.5, respectively; $P = 0.002$), examinations with a reference standard PI-RADS score of 4 (from 5.9 ± 2.4-6.8 ± 2.4, respectively; $P = 0.032$), and examinations with a reference standard PI-RADS score of 5 (from 7.5 ± 2.6-7.9 ± 2.1, respectively; $P = 0.024$). There was no significant difference in confidence for examinations positive for transition zone tumor (from 6.6 ± 2.2-6.0 ± 2.2, respectively; $P = 0.398$) or examinations with a reference standard PI-RADS score of 3 (from 5.0 ± 2.5-5.4 ± 2.3, respectively; $P = 0.857$).

There was no significant difference between sessions 1 and 2 for the the overall concordance among all possible pairs of readers in interpreting examinations as positive or negative (from 81.7% [147/180]-76.7% [138/180], respectively; $P = 0.230$). However, there was a significant improvement in session 2 for the overall concordance among all possible pairs of readers in correctly detecting tumors (from 61.1% [55/90]-73.3% [66/90], respectively; $P = 0.027$]. Among 50 instances of a given pair of readers correctly detecting a tumor during both sessions, there was no significant difference between sessions for concordance of the provided PI-RADS score (from 64.0% [32/50]-70.0% [35/50], respectively; $P = 0.607$). Kappa coefficients for interpretation of examinations as positive or negative indicated moderate reproducibility (0.53) in session 1 compared with substantial reproducibility (0.63) in session 2. Kappa coefficients for PI-RADS scores for tumors correctly detected by a given pair of readers during both sessions indicated moderate reproducibility in both sessions (0.56 in session 1 and 0.57 in session 2).

Table 5 summarizes results of the survey completed before and after website review. Small improvements were noted in readers' reported comfort in interpreting prostate MRI, confidence in ability to detect lesions, and confidence in ability to score lesions using PI-RADS v2. The 3 readers all reported that the interactive website took from 3.5-4 hours to review. All 3 readers indicated that the number of available cases, the overall level of difficulty, and the amount of time required to review the cases, were "about right." Following website review, the readers most strongly agreed (maximal level of agreement of 5 provided by all 3 readers) that the cases were useful as a learning tool, that they were likely to view the cases again, that the interactive format enhanced the cases' educational value, and that review of the cases improved their skill in prostate MRI interpretation. The readers had lowest levels of agreement (average level of agreement of 3.7 on 1-5 scale) that the cases addressed a satisfactory range of diagnoses and that the explanations of the interactive cases were satisfactory.

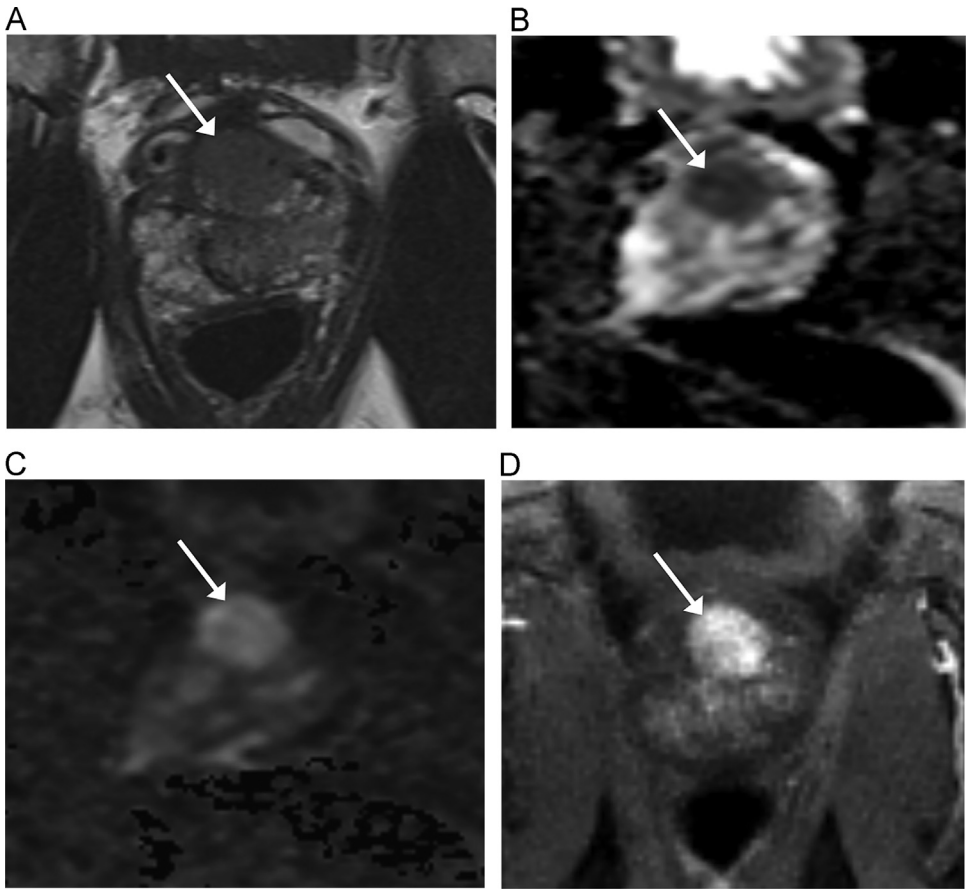

FIG 4. A 56 year-old man with Gleason score 3 + 4 tumor on MRI-ultrasound fusion-targeted biopsy of midline anterior transition zone lesion. The lesion was correctly detected by 2 of the 3 readers during both sessions, and missed by the remaining reader during both sessions. For the 2 readers who detected the lesion during both sessions, confidence in the interpretation increased in session 2 following review of the online interactive case material.

Discussion

The online interactive case material explored in this investigation resulted in significant improvements in readers' sensitivity and NPV for tumor detection on prostate MRI. This observation may relate to the feedback that readers received while reviewing the online cases, being informed of the presence and location of tumors that they missed during their initial blinded review. The improvements in detection were most notable for lesions with a relatively lower PI-RADS score and a lower Gleason score, suggesting particular value for helping novice readers of prostate MRI in detecting potentially more subtle or challenging lesions. The interactive online material also resulted in widespread significant improvements in reader confidence, for both positive and negative examinations. Such improvement in reader confidence is helpful for fostering more firmly crafted reports that may provide clearer guidance to referring physicians in clinical decision-making.

The improved sensitivity as a result of the online material was not accompanied by significant improvements in specificity or PPV. Incorporating into the online case material a larger volume of normal examinations, and of examinations with benign findings that mimic tumor, may help the readers reduce the frequency of false-positive interpretations. In addition, the significantly improved sensitivity and reader confidence was observed solely for tumors in the peripheral zone, and not for those in the transition zone. A larger number of examples transition zone tumors among the online cases may, therefore, have been helpful as well. Detection of transition zone tumors is recognized to be particularly challenging,[35,36] and an earlier study reported that a targeted didactic reader education program had the largest diagnostic impact for anterior cancers.[25] Thus, it remains important for prostate MRI educational initiatives to specifically address this aspect of interpretation.

Despite the improvements in sensitivity and reader confidence, the online material did not lead to a comparable improvement in correct assignments of PI-RADS scores to tumors. The annotated solutions to the online material provided a PI-RADS score for each examination, as well as a brief explanation. However, PI-RADS remains a subjective system based on qualitative visual features, without strict quantitative criteria for selecting a given score. In addition, correctly identifying the exact PI-RADS score for a given case presents an additional level of complexity beyond just tumor detection in a binary fashion (ie, tumor present vs absent) given the need to select a more granular score from among a multipoint scale. Achieving improved PI-RADS scoring may require viewing a larger volume of cases than included in the online material for this investigation, more intensive didactic instruction regarding the nuances of the system, or feedback of a more individualized nature beyond simply providing a solution but also explaining why a reader's suggested PI-RADS score was incorrect.

Improving overall community performance in prostate MRI interpretation currently represents a pressing issue within the field.[22,23] To address this issue, numerous national radiology specialty societies have sponsored in-person hands-on prostate MRI educational courses[37-39] that incorporate case review with feedback from a panel of experienced instructors. However, such traditional educational initiatives, beyond high costs and logistical challenges from the reliance upon on-site hardware for case access, have inherent limitations in impact. The courses provide only a single exposure of finite duration for a limited number of

TABLE 2
Comparison of sensitivity between sessions 1 and 2 for subsets of examinations based on tumor characteristics

Feature	Reader	Session 1		Session 2		P^*
		%	*n*	%	*n*	
PI-RADS 3 lesion by reference standard	1	14.3	1/7	42.9	3/7	0.500
	2	14.3	1/7	42.9	3/7	0.500
	3	14.3	1/7	42.9	3/7	0.625
	All	14.3	3/21	42.9	9/21	0.070
PI-RADS 4 lesion by reference standard	1	58.3	7/12	91.7	11/12	0.125
	2	75.0	9/12	75.0	9/12	1.000
	3	50.0	6/12/	66.7	8/12	0.500
	All	61.1	22/36	77.8	28/36	0.070
PI-RADS 5 lesion by reference standard	1	90.9	10/11	100.0	11/11	1.000
	2	81.8	9/11	90.9	10/11	1.000
	3	72.7	8/11	72.7	8/11	1.000
	All	81.8	27/33	87.9	29/33	0.625
GS 3 + 4 tumor	1	53.3	8/15	86.7	13/15	0.063
	2	53.3	8/15	60.0	9/15	1.000
	3	40.0	6/15	66.7	10/15	0.125
	All	48.9	22/45	71.1	32/45	**0.006**
GS ≥ 4 + 3 tumor	1	66.7	10/15	80.0	12/15	0.500
	2	73.3	11/15	86.7	13/15	0.500
	3	60.0	9/15	60.0	9/15	1.000
	All	66.7	30/45	75.6	34/45	0.289
Peripheral zone tumor	1	54.5	12/22	81.8	18/22	**0.031**
	2	72.7	16/22	90.9	20/22	0.125
	3	54.5	12/22	68.2	15/22	0.375
	All	60.6	40/66	80.3	53/66	< **0.001**
Transition zone tumor	1	75.0	6/8	87.5	7/8	1.000
	2	37.5	3/8	25.0	2/8	1.000
	3	37.5	3/8	50.0	4/8	1.000
	All	50.0	12/24	54.2	13/24	1.000

* Highlighted in bold when statistically significant at $P < 0.05$.

TABLE 3
Comparison of percentage of examinations with an accurate PI-RADS score between sessions

Zone	Reader	Session 1		Session 2		*P* value
		%	*n*	%	*n*	
All tumors						
Both	1	33.3	10/30	56.7	17/30	0.092
Both	2	43.3	13/30	53.3	16/30	0.453
Both	3	53.3	16/30	50.0	15/30	1.000
Both	All	43.3	39/90	53.3	48/90	0.149
Peripheral	1	36.4	8/22	59.1	1/22	0.180
Peripheral	2	45.5	10/22	63.6	14/22	0.125
Peripheral	3	54.5	12/22	54.5	12/22	1.000
Peripheral	All	45.5	30/66	59.1	39/66	0.078
Transition	1	25.0	2/8	50.0	4/8	0.625
Transition	2	37.5	3/8	25.0	2/8	1.000
Transition	3	50.0	4/8	37.5	3/8	1.000
Transition	All	37.5	9/24	37.5	9/24	1.000
*Tumor correctly detected by given reader in both sessions**						
Both	1	50.0	9/18	66.7	12/18	0.453
Both	2	60.0	12/20	65.0	13/20	1.000
Both	3	73.7	14/19	68.4	13/19	1.000
Both	All	61.4	35/57	66.7	38/57	0.629
Peripheral	All	62.8	27/43	69.8	30/43	0.508
Transition	All	57.1	8/14	57.1	8/14	1.000

* The subset of tumors correctly detected by a given reader during both sessions was not large enough to allow results to be stratified by both zone and by reader; thus, for these examinations, results are only stratified either by zone or by reader.

participants who are able to attend. Also, the timing of the course may not match when a given learner desires to access the material. Moreover, a wider variety of audiences may be interested in the course material than typically attend the traditional in-person courses. Although individual departments may seek to create local educational programs, departments' resource and expertise in the topic is variable. In addition, it is inefficient for multiple departments to prepare similar didactic content on the same topic, as well as to offer similar content on multiple occasions.

In comparison, the online interactive case material allows learners to drive the accessing of the material when it is most relevant to them, unrestricted for time or location. Learners may readily review selected portions of the content that are of greatest personal interest, including revisiting material on multiple occasions at their own discretion. Not only can a larger volume of material be provided, but the online content is available to a considerably larger number of learners, including more diverse audiences[40] that encompass radiologists of multiple levels of experience, referring physicians, and patients. In this regard, the online material allows for a more flexible educational experience that can be tailored to meet learners' individual goals. Moreover, the online content provides standardized material that can be used in a consistent fashion across programs nationally regardless of their size, resources, and expertise in the topic, potentially providing exposure to the material for learners that otherwise may have lesser opportunity to receive robust instruction in the area.

The survey indicated that the 3 readers responded favorably to the online material and its perceived educational value, reporting improved levels of confidence and comfort in prostate MRI

TABLE 4
Comparison of reader confidence between sessions

Cases	Reader	Session 1	Session 2	P^*
All	1	3.2 ± 1.9	3.6 ± 2.0	0.097
	2	7.0 ± 2.4	7.5 ± 1.4	0.244
	3	6.5 ± 1.8	7.9 ± 1.5	**< 0.001**
	All	5.6 ± 2.7	6.3 ± 2.6	**0.026**
Benign	1	2.6 ± 1.6	2.6 ± 1.3	0.848
	2	6.5 ± 2.5	7.1 ± 1.2	0.251
	3	5.6 ± 1.1	7.4 ± 1.2	**< 0.001**
	All	4.9 ± 2.5	5.7 ± 2.5	**0.004**
Tumor	1	3.9 ± 2.0	4.6 ± 2.1	**0.046**
	2	7.6 ± 2.1	7.8 ± 1.6	0.560
	3	7.4 ± 2.0	8.4 ± 1.7	**0.013**
	All	6.3 ± 2.7	6.9 ± 2.4	**0.009**
PI-RADS 3 lesion by reference standard	1	2.6 ± 1.8	2.6 ± 0.8	1.0
	2	6.7 ± 2.1	6.4 ± 1.0	0.753
	3	5.7 ± 1.3	7.3 ± 1.3	0.106
	All	5.0 ± 2.5	5.4 ± 2.3	0.857
PI-RADS 4 lesion by reference standard	1	3.5 ± 1.8	4.2 ± 1.6	0.314
	2	7.2 ± 2.1	7.9 ± 1.2	0.205
	3	7.1 ± 1.2	8.3 ± 1.8	0.062
	All	5.9 ± 2.4	6.8 ± 2.4	**0.032**
PI-RADS 5 lesion by reference standard	1	5.1 ± 1.8	6.3 ± 1.9	**0.009**
	2	8.6 ± 1.9	8.5 ± 1.8	1.0
	3	8.8 ± 2.1	9.1 ± 1.4	1.0
	All	7.5 ± 2.6	7.9 ± 2.1	**0.024**
GS 3 + 4 tumor	1	3.5 ± 1.9	4.0 ± 1.5	0.102
	2	7.5 ± 2.2	7.6 ± 1.6	0.249
	3	6.9 ± 1.8	8.7 ± 1.2	**< 0.001**
	All	6.0 ± 2.6	6.8 ± 2.5	**0.026**
GS ≥ 4 + 3 tumor	1	4.2 ± 2.1	5.1 ± 2.5	**0.014**
	2	7.7 ± 2.1	7.9 ± 1.6	0.469
	3	7.9 ± 2.1	8.1 ± 2.0	0.797
Peripheral zone tumor	1	3.6 ± 2.0	4.5 ± 2.2	**0.008**
	2	8.1 ± 2.0	8.2 ± 1.5	0.643
	3	7.6 ± 2.0	8.3 ± 1.8	**0.045**
	All	6.4 ± 2.8	7.0 ± 2.5	**0.002**
Transition zone tumor	1	4.6 ± 2.1	4.8 ± 2.4	1.000
	2	6.4 ± 2.1	6.6 ± 1.2	0.844
	3	7.0 ± 2.1	8.5 ± 1.1	0.219
	All	6.0 ± 2.2	6.6 ± 2.2	0.398

* Highlighted in bold when statistically significant at $P < 0.05$.

interpretation. The overall volume of material, as well as its level of difficulty, were deemed appropriate, and the readers all completed the material within 4 hours. Given the observations, the online interactive material may provide an enduring educational resource of value not only for interpreting radiologists, but referring physicians as well. Ultimately, it is intended that the educational value of the website can translate to improved care for prostate cancer patients. In addition, the interactive online material may serve as a model to apply for development of similar educational content for other areas of imaging.

This study has a number of limitations. First, assessment was only performed among 3 readers from a single institution. Greater variability in results may have been observed among a larger number of readers, including readers from different institutions. In addition, the effect of the online material was only assessed for learners who were novices in prostate MRI interpretation. The potential value of the material for more experienced readers remains unknown. Also, it remains unknown from the analysis to what extent the observed improvements resulted from the annotated solutions provided by the online material vs the inherent exposure to a larger number of examinations while viewing the material. Finally, while the reference standard for evaluating the readers' assigned PI-RADS scores was established by a radiologist experienced in prostate MRI and PI-RADS v2, these scores remain subjective, with no perfect measure of ground "truth."

In conclusion, an online interactive case-based website in prostate MRI interpretation could be completed within 4 hours and resulted in significantly improved sensitivity and NPV for tumor detection, along with significantly improved reader confidence, among novice readers of prostate MRI. This online material may serve as a resource that complements existing traditional methods of instruction by allowing for a more flexible educational experience among a larger volume of learners. However, the online material did not result in improved PI-RADS scoring of detected tumors, such that further more targeted educational initiatives regarding the application of PI-RADS v2 are warranted.

TABLE 5
Results of survey

Survey item	Response*
Items included on survey completed both before and after website review	
I feel comfortable interpreting prostate MRI examinations. (1-5)	Pre: 2.3; Post: 3.7
I am confident in my ability to detect lesions. (1-5)	Pre: 2.7; Post: 3.7
I am confident in my ability to score lesions using PI-RADS version 2. (1-5)	Pre: 2.0; Post: 3.3
Items included on survey completed only after website review	
How long did it take to complete to complete the cases? (free response)	3.5 hours ($n = 1$); 4 hours ($n = 2$)
The number of available cases on this website is: (multiple choice)	"About right" ($n = 3$)
The overall level of difficulty of the cases on this website is: (multiple choice)	"About right" ($n = 3$)
The amount of time required to review the cases on this website is: (multiple choice)	"About right" ($n = 3$)
I am likely to view the interactive cases again. (1-5)	5.0
The interactive cases are useful as a learning tool. (1-5)	5.0
The interactive cases had a user-friendly interface. (1-5)	4.7
The interactive format of the cases enhances their educational value. (1-5)	5.0
The interactive cases satisfactorily addressed my questions on this topic. (1-5)	4.0
The interactive cases addressed a satisfactory range of diagnoses (1-5)	3.7
The explanations of the interactive cases are satisfactory. (1-5)	3.7
Reviewing the interactive cases improved my skill in prostate MRI interpretation. (1-5)	5.0
Reviewing the interactive cases improved my knowledge of PI-RADS. (1-5)	4.7

* For all responses on a 1-5 scale, 5 indicates maximal agreement, and provided value indicates mean response from among the 3 readers.

References

1. Hoeks CM, Barentsz JO, Hambrock T, et al. Prostate cancer: Multiparametric MR imaging for detection, localization, and staging. Radiology 2011;261(1):46–66.
2. Lee DJ, Ahmed HU, Moore CM, et al. Multiparametric magnetic resonance imaging in the management and diagnosis of prostate cancer: current applications and strategies. Curr Urol Rep 2014;15(3):390.
3. de Rooij M, Crienen S, Witjes JA, et al. Cost-effectiveness of magnetic resonance (MR) imaging and MR-guided targeted biopsy versus systematic transrectal ultrasound-guided biopsy in diagnosing prostate cancer: A modelling study from a health care perspective. Eur Urol 2014;66(3):430–6.
4. Lotan Y, Haddad AQ, Costa DN, et al. Decision analysis model comparing cost of multiparametric magnetic resonance imaging vs repeat biopsy for detection of prostate cancer in men with prior negative findings on biopsy. Urol Oncol 2015;33(6):266.e9-16.
5. Rosenkrantz AB, Verma S, Choyke P, et al. Prostate magnetic resonance imaging and magnetic resonance imaging targeted biopsy in patients with a prior negative biopsy: A consensus statement by AUA and SAR. J Urol 2016;196(6):1613–8.
6. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology. Prostate Cancer Early Detection. Version 2.2015. http://www.nccn.org/professionals/physician_gls/PDF/prostate_detection.pdf.
7. Heidenreich A, Bastian PJ, Bellmunt J, et al. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. Eur Urol 2014;65(1):124–37.
8. American Urological Association. MR US Fusion Guided Prostate Biopsy. https://university.auanet.org/liveCourse-ProductInformation.cfm?id=3166. [Accessed on August 28, 2017].
9. Barentsz JO, Weinreb JC, Verma S, et al. Synopsis of the PI-RADS v2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. Eur Urol 2016;69(1):41–9.
10. Westphalen AC, Rosenkrantz AB. Prostate imaging reporting and data system (PI-RADS): Reflections on early experience with a standardized interpretation scheme for multiparametric prostate MRI. AJR Am J Roentgenol 2014;202(1):121–3.
11. Rosenkrantz AB, Taneja SS. Radiologist, be aware: Ten pitfalls that confound the interpretation of multiparametric prostate MRI. AJR Am J Roentgenol 2014;202(1):109–20.
12. Rosenkrantz AB, Verma S, Turkbey B. Protesta cancer: Top places where tumors hide on multiparametric MRI. AJR Am J Roentgenol 2015;204(4):W449–56.
13. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS prostate imaging—Reporting and data system: 2015, Version 2. Eur Urol 2016;69(1):16–40.
14. American College of Radiology. Prostate Imaging Reporting and Data System (PI-RADS). http://www.acr.org/Quality-Safety/Resources/PIRADS. [Accessed August 28, 2017].
15. Barrett T, Turkbey B, Choyke PL. PI-RADS version 2: What you need to know. Clin Radiol 2015;70(11):1165–76.
16. Wahab SA, Verma S. Review of prostate imaging reporting and data system version 2. Future Oncol 2016;12(21):2479–94.
17. Spektor M, Mathur M, Weinreb JC. Standards for MRI reporting-the evolution to PI-RADS v 2.0. Transl Androl Urol 2017;6(3):355–67.
18. Shaish H, Taneja SS, Rosenkrantz AB. Prostate MR imaging: An update. Radiol Clin North Am 2017;55(2):303–20.
19. Horn GL Jr., Hahn PF, Tabatabaei S, et al. A practical primer on PI-RADS version 2: A pictorial essay. Abdom Radiol (NY) 2016;41(5):899–906.
20. Franiel T, Asbach P, Teichgraber U, et al. Prostate imaging—An update. Rofo 2015;187(9):751–9.
21. 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(3):793–804.
22. Gupta RT, Spilseth B, Froemming AT. How and why a generation of radiologists must be trained to accurately interpret prostate mpMRI. Abdom Radiol (NY) 2016;41(5):803–4.
23. Puech P, Randazzo M, Ouzzane A, et al. How are we going to train a generation of radiologists (and urologists) to read prostate MRI? Curr Opin Urol 2015;25(6):522–35.
24. Akin O, Riedl CC, Ishill NM, et al. Interactive dedicated training curriculum improves accuracy in the interpretation of MR imaging of prostate cancer. Eur Radiol 2010;20(4):995–1002.
25. Garcia-Reyes K, Passoni NM, Palmeri ML, 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(1):134–42.
26. Papamichail D, Pantelis E, Papagiannis P, et al. A Web simulation of medical image reconstruction and processing as an educational tool. J Digit Imaging 2015;28(1):24–31.
27. Meckfessel S, Stuhmer C, Bormann KH, et al. Introduction of e-learning in dental radiology reveals significantly improved results in final examination. J Craniomaxillofac Surg 2011;39(1):40–8.
28. Johnston J. Effectiveness of online instruction in the radiologic sciences. Radiol Technol 2008;79(6):497–506.
29. Pinto A, Selvaggi S, Sicignano G, et al. E-learning tools for education: regulatory aspects, current applications in radiology and future prospects. Radiol Med 2008;113(1):144–57.
30. Learn Prostate MRI. Interactive Cases. http://learnprostatemri.com. [Accessed August 28, 2017].
31. Rosenkrantz AB, Ayoola A, Hoffman D, et al. The LearningCurve in Prostate MRI Interpretation: Self-Directed Learning Versus ContinualReader Feedback. AJR Am J Roentgenol 2017;208(3):W92–100.
32. Rosenkrantz AB, Babb JS, Taneja SS, Ream JM. Proposed Adjustmentsto PI-RADS Version 2 Decision Rules: Impact on Prostate Cancer Detection. Radiology 2017;283(1):119–29.
33. Mendhiratta N, Rosenkrantz AB, Meng X, et al. MagneticResonance Imaging-Ultrasound Fusion Targeted Prostate Biopsy in a Consecutive Cohort of Men with No Previous Biopsy: Reduction of Over Detection throughImproved Risk Stratification. J Urol 2015;194(6):1601–6.
34. Meng X, Rosenkrantz AB, Mendhiratta N, et al. Relationship Between Prebiopsy Multiparametric Magnetic Resonance Imaging (MRI), Biopsy Indication, and MRI-ultrasound Fusion-targeted Prostate Biopsy Outcomes. Eur Urol 2016;69(3):512–7.
35. Hoeks CM, Hambrock T, Yakar D, et al. Transition zone prostate cancer: Detection and localization with 3-T multiparametric MR imaging. Radiology 2013;266(1):207–17.

36. Rosenkrantz AB, Kim S, Campbell N, et al. Transition zone prostate cancer: Revisiting the role of multiparametric MRI at 3 T. AJR Am J Roentgenol 2015;204(3):W266–72.
37. American Roentgen Ray Society. Event Registration—2016 ARRS Prostate MR Imaging Symposium. http://store.arrs.org/events/registration.aspx?event=PROSTATE16. [Accessed August 28, 2017].
38. Radiological Society of North America. RSNA 2016 Daily Bulletin. New Hands-on Prostate MRI Course is a Hit. https://rsna2016.rsna.org/dailybulletin/index.cfm?pg=16tue78. Published November 29, 2016; [Accessed August 28, 2017].
39. Society of Abdominal Radiology. Program Highlights. https://www.abdominalradiology.org/?page=ProgramHighlights. [Accessed August 28, 2017].
40. Bhargava P, Dhand S, Lackey AE, et al. Radiology education 2.0—On the cusp of change: Part 2. eBooks; file sharing and synchronization tools; websites/teaching files; reference management tools and note taking applications. Acad Radiol 2013;20(3):373–81.