Prostate Imaging Reporting and Data System (PI-RADS): Reflections on Early Experience With a Standardized Interpretation Scheme for Multiparametric Prostate MRI

Antonio C. Westphalen¹
Andrew B. Rosenkrantz²

OBJECTIVE. In this article, we provide our perspective on current multiinstitutional efforts to improve standardization of interpretation and reporting of multiparametric prostate MRI, emphasizing their strengths and limitations based on our experiences, and provide several suggestions for guiding continued initiatives for standardizing multiparametric prostate MRI reporting.

CONCLUSION. Significant steps are being taken to facilitate the adoption of prostate MRI by urologists and other physicians in the community. Ultimately, however, prospective multicenter validation studies assessing the various aspects of a diagnostic test will be required.

Although there have been major improvements in the diagnostic performance and technical standardization of multiparametric prostate MRI in the past decade, widespread clinical adoption remains hindered by the lack of standardization of interpretation and reporting. Nonetheless, significant multiinstitutional efforts have attempted to address this challenge in recent years [1–3]. In this article, we provide our perspective on such efforts, emphasizing their strengths and limitations based on our experiences, and provide several suggestions for guiding continued initiatives for standardizing multiparametric prostate MRI reporting.

Medical diagnostic tests must be accurate and precise. Accuracy, in this context, refers to how close a measurement is from the actual true value, whereas precision is the reproducibility or repeatability of such results under stable conditions. Good accuracy and precision have, in general, provided the basis for the widespread acceptance and clinical use of many time-tested examinations. Moreover, standardization allows robust investigation of medical tests through well-designed research to ultimately determine whether such tests result in improved patients’ outcomes. Although many single-center studies have shown the value of MRI for the diagnosis and staging of prostate cancer, prostate MRI has not yet reached the same acceptance level of many other medical tests. This variability in expectations among urologists regarding the role of prostate MRI relates, in part, to the wide range of treatment options available for a given stage of disease as well as a potential lack of local radiologic expertise. However, a more important reason for the lack of adoption of prostate MRI into clinical practice may be a lack of standardization among radiologists, including those experienced in prostate imaging.

Although prostate MRI acquisition protocols have become much more uniform in recent years, there is an urgent need for a standard approach to interpret and report (i.e., formally communicate) the MRI findings.

Several different interpretative approaches for multiparametric MRI, of varying level of complexity, have been suggested in the literature. Pinto et al. [4], for example, have proposed a scheme in which each MRI parameter is categorized as positive or negative in a binary fashion; according to the number of sequences that are categorized as positive, an overall score of low, intermediate, or high is assigned to each lesion. Although a strength of this method is its straightforward approach, a limitation is that, by considering each sequence as simply positive or negative, it fails to account for variation in the degree to which a given sequence is abnormal, which can be substantial and may affect the potential significance of the lesion. Also, this system treats tumors in the peripheral and transition zone the same, although prior studies have shown that the diagnostic performance of a given parameter changes depending on the location of the tumor [5].

Keywords: interpretation, MRI, prostate cancer, reporting

DOI:10.2214/AJR.13.10889
Received March 10, 2013; accepted after revision April 23, 2013.

¹Department of Radiology and Biomedical Imaging, University of California San Francisco, 505 Parnassus Ave, M-372, Box 0628, San Francisco, CA, 94143. Address correspondence to A. C. Westphalen (antonio.westphalen@ucsf.edu).
²Department of Radiology, New York University, New York, NY.

AJR 2014; 202:121–123
0361–803X/14/2021–121
© American Roentgen Ray Society
Two recent large multiinstitutional efforts have attempted to address these problems related to standardization of imaging interpretation and reporting [2, 3]. Dickinson et al. [3] recently published the recommendations of a consensus meeting of 16 European urologists and radiologists with expertise in prostate cancer, and the European Society of Urogenital Radiology published the initial description of a structured reporting system called PI-RADS (Prostate Imaging Reporting and Data System) [2]. Although both panels deserve congratulations for their efforts, it must be understood that these represent works in progress and that further refinement will be required for either approach to achieve maximal impact.

The publication of Dickinson et al. [3] sheds important light into the needs and expectations that urologists and other clinicians have of radiologists who interpret prostate MRI. However, the scheme essentially states that cases should be reported on a 1–5 scale without suggesting any criteria or providing any assistance in how to interpret images and derive such a score. Thus, this work addresses only a small part of the problem and does not provide a complete solution for standardization.

PI-RADS is a new, but already well-known, scheme that offers a major advance by providing an algorithm for interpreting multiparametric MRI of the prostate [2]. This classification was initially developed with the success of BI-RADS in mind, which is used worldwide and has led to significant advances in breast cancer diagnosis and treatment, as well as the widely standardized implementation of screening mammography for cancer detection. It is expected that similar results could be seen in prostate cancer with the use of MRI. Nonetheless, a direct comparison between PI-RADS and BI-RADS is difficult, because treatment of breast cancer is much more standardized than treatment of prostate cancer. Among the reasons for the early notoriety and adoption of PI-RADS, in spite of its still being an evolving tool, is the fact that it was developed and is supported by some of the most well-known institutions in the field, including the European Society of Urogenital Radiology, the AdMeTech Foundation, and the American College of Radiology [1].

As with the proposal by Dickinson et al. [3], PI-RADS states that each lesion should be described using a 1–5 scale. Specifically, PI-RADS states that each lesion should receive a score from 1 to 5 for each individual imaging parameter, as well as an overall 1–5 score reflecting the likelihood of clinically significant cancer. Importantly, PI-RADS goes beyond and provides a short verbal description of findings that correspond to each score from 1 through 5 for each individual sequence, thereby taking a major step forward in addressing the issue of standardized interpretation.

In our opinion, the primary limitation of PI-RADS in its current form is that the original publication fails to explain how the overall 1–5 score is derived. Although it may be intuitive to use a “sum score” (i.e., the sum of the individual 1–5 scores provided for each sequence), this is not necessarily the best option. First, the use of an overall 1–5 scale is widespread in reporting schemes for diagnostic imaging studies, as evidenced by both BI-RADS and LI-RADS (Liver Imaging Reporting and Data System), and has gained familiarity by radiologists and clinicians alike; it is thus desirable to parallel this approach with PI-RADS. In addition, according to PI-RADS, T2-weighted MRI, diffusion-weighted MRI, and dynamic contrast-enhanced MRI should all be routinely acquired, whereas MR spectroscopic imaging is considered optional. Therefore, the use of a sum score would lead to lesions being scored on either a 3–15 or 4–20 scale, depending on the use of MR spectroscopic imaging at a given center. This variability in the range of the scale would confound multiinstitutional efforts to standardize and determine the exact relationship between the study results and likelihood of cancer or cancer aggressiveness in a given patient. Accordingly, a single overall 5-point score that takes into account the scores of all MRI parameters may be a more appropriate way of conveying the results of MRI using PI-RADS. In addition, recent preliminary data suggest that such an approach has at least moderate interreader agreement [6]. Finally, there is currently a lack of consensus among urologists and pathologists as to the histologic criteria for “clinically significant” cancer. Although the PI-RADS score from 1 to 5 is intended to indicate the likelihood of clinically significant cancer, the publication does not provide a precise definition of cancers that are deemed significant and, therefore, being evaluated using the proposed system.

Although some may suggest that routinely reporting scores for each MRI parameter, in addition to the overall 1–5 score, is not necessary and may be confusing, we think that inclusion of these scores can add value by enhancing the understanding of the final interpretation by clinicians familiar with the individual parameters, aiding research efforts comparing the clinical significance of the sequences, allowing more robust interinstitutional comparisons, and assisting the training of less-experienced radiologists. We suggest that a simple way of lessening any confusion related to the possible role of the previously noted sum score is to replace the 1–5 designations for the individual sequences with letters. That is, one could classify findings from A to E for each sequence, rather than applying a series of 1–5 scores. This alternative avoids the psychologic urge to add up the numbers, while still preserving all the information in the report.

An additional concern is that the current PI-RADS publication does not indicate how to weigh the scores given to the individual sequences. Previous studies have shown that the diagnostic and prognostic values of each MRI parameter are not the same and, in fact, change depending on the location of the lesion in the peripheral or in the transition zone. For instance, diffusion-weighted MRI has been observed to be the best-performing single parameter in the peripheral zone [7], suggesting that it may deserve the strongest weighting in the peripheral zone. On the other hand, T2-weighted MRI has been shown to have a critical role for identification of transition zone tumors given its depiction of the morphologic features and texture of lesions [8] and may, therefore, warrant the strongest weighting in the transition zone. Furthermore, dynamic contrast-enhanced MRI is not reliable for diagnosing prostate cancer in the transition zone, given the significant overlap with benign prostatic hyperplasia [9]. Thus, it is possible that this technique should be essentially discounted when evaluating transition zone lesions. Any such variable weighting of sequences in the different zones requires a balance between simplicity of use and efficacy.

A last potential source of confusion is the definition of PI-RADS score of 3 for T2- and diffusion-weighted MRI. As currently stated, a finding receives a score of 3 for these sequences if it does not meet the criteria for any of the other scores. Although this approach may increase the specificity of findings classified as 1, 2, 4, and 5, it raises the possibility that a heterogeneous spectrum of abnormalities will all be scored identically as 3. Thus, more precise definition of findings comprising a score of 3 may be warranted.

An overall suggestion that we provide, to address a number of these limitations, is to establish a flowchart that converts all possible combinations of the individual sequence classifications to a final overall 1–5 score, thereby allowing separate weighting of the
PI-RADS and Multiparametric Prostate MRI

![Diagram](https://via.placeholder.com/150)

**Fig. 1**—Example of approach for reporting multiparametric prostate MRI that incorporates graphical and tabular elements, as based on suggestions provided in this article. Prostate Imaging Reporting and Data System (PI-RADS) score of 4 reflects heavier weighting of findings on diffusion-weighted MRI (DWI) relative to other sequences for peripheral zone lesions. Note that complete report would include additional elements, such as patient demographics, detailed clinical information, prostate volume, and other findings related to tumor staging. a = anterior, as = anterior stroma, DCE = dynamic contrast-enhanced MRI, EPE = extraprostatic extension, L = left, MRSI = MR spectroscopic imaging, NA = not applicable, p = posterior, R = right, SV = seminal vesicle, T2WI = T2-weighted MRI. Adapted with permission from [3].

Table 1

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Size (mm)</th>
<th>T2WI (A–E)</th>
<th>DWI (A–E)</th>
<th>DCE (A–E)</th>
<th>MRSI (A–E)</th>
<th>PI-RADS Score (1–5)</th>
<th>EPE (Y/N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12 × 8</td>
<td>C</td>
<td>D</td>
<td>C</td>
<td>n/a</td>
<td>4</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

individual sequences and different methods for conversion to the overall score in the peripheral and transition zones. This could provide an elegant and easy way of conveying the distinct roles of the individual sequences and the overall scores, while keeping PI-RADS simple enough for everyday use.

One last consideration is the standardization of the actual structure and format of the final report, which can be aided by use of the PI-RADS system. The PI-RADS publication calls for localizing abnormalities using at least a 16-region, and optimally a 27-region, scheme. Such precise localization is important for guiding targeted biopsy, focal therapy, and surveillance. However, this can be difficult for a radiologist to describe in words and for a urologist to understand. Thus, software needs to be developed to allow radiologists to generate graphical representations depicting the location of reported lesions, in a manner that is integrated into routine clinical workflow. Indeed, some centers have already achieved such solutions. In Figure 1, we provide one example of many possible alternatives to convey the final interpretation.

In conclusion, we applaud the radiology community for its collaborative efforts, because standardization is a critical step toward accelerating the adoption of prostate MRI by urologists and other physicians in the community. We hope our suggestions will prove useful in these efforts. Ultimately, however, prospective multicenter validation studies assessing the various aspects of a diagnostic test, such as intra- and interreader variability, diagnostic accuracy, discriminatory ability, and predictive value, will be required.

**References**


This article has been cited by:


2. Jeffrey S. Quon, Bardia Moosavi, Maneesh Khanna, Trevor A. Flood, Christopher S. Lim, Nicola Schieda. 2015. False positive and false negative diagnoses of prostate cancer at multi-parametric prostate MRI in active surveillance. *Insights into Imaging*. [CrossRef]


