Radiology: Imaging Cancer

COMMENTARY

A Grading System for Extraprostatic Extension of Prostate Cancer That We Can All Agree Upon?

Peter L. Choyke, MD

Peter L. Choyke, MD, is a senior investigator and chief of the Molecular Imaging Program of the National Cancer Institute. He received his medical degree from Jefferson Medical College and trained in diagnostic radiology at Yale and completed a fellowship at Penn. He has focused on the imaging of genitourinary malignancies, particularly prostate cancer. Along with his colleagues Brad Wood and Peter Pinto, Dr Choyke developed the first MRI/US fusion biopsy system for prostate cancer. He served on the PI-RADS development committee for versions 1 and 2. He is currently a faculty member on the American College of Radiology's prostate MRI course. His current research interests include AI of prostate cancer imaging and pathology, image-guided therapies, augmented reality, and prostate-specific membrane antigen PET scanning of prostate cancer.

Prostate cancer exhibits a wide range of aggressiveness from indolent to lethal biologic behavior. Recognizing this, treatments become more radical as the number of adverse features associated with the cancer increase, although with more radical treatments come increasing side effects. Thus, to direct optimal, individualized treatment, it is important to characterize the disease status as accurately and consistently as possible.

Extraprostatic extension (EPE) is an important indicator of prostate cancer aggressiveness and is an independent predictor of biochemical recurrence-free survival, which is strongly correlated with the histologic severity of EPE as measured by radial extension of the cancer from the gland (1–4). Because of the higher risk of recurrence, more aggressive therapies are warranted when EPE is clearly present. Because EPE is difficult to determine intraoperatively, the diagnosis must be made preoperatively. Parenthetically, *EPE* is the preferred term as the prostate does not have a well-formed capsule and therefore *extracapsular extension* is less appropriate. Thus, EPE is associated with a higher risk of recurrence, metastases, and ultimately death and is therefore assigned importance by clinicians.

The Prostate Imaging Reporting and Data System (PI-RADS) has been instrumental in improving the uniformity of reporting prostate cancer. It provides standardized nomenclature and criteria with which to assess the risk that a lesion identified at multiparametric MRI represents clinically significant cancer. The establishment of PI-RADS has had many desirable effects. First, it provides a standardized basis for teaching radiologists how to interpret multiparametric MRI. Additionally, it provides a shorthand for communication among the various subspecialists who treat prostate cancer, thus facilitating accurate conveyance of important clinical information. It also provides a basis with which to compare studies in the literature, enabling more meaningful meta-analyses.

While PI-RADS addresses EPE, it does so in a vague way, mentioning the various features such as tumor contact length, irregularity, bulging, gross extension, and loss of rectoprostatic angle, but it does not assign relative predictive values to these features. (For reference, contact length is the perceived distance that the tumor contacts the outer margin of the prostate.) To address the shortcoming of PI-RADS, Mehralivand et al proposed a three-point grading system for EPE that predicts its likelihood at histologic examination based on experience with more than 500 prospectively recruited patients undergoing multiparametric MRI prior to surgery (5). Grade 1 refers to tumors with a contact length of 1.5 cm or greater or contour bulge or irregularity. Grade 2 refers to tumors with a contact length of 1.5 cm or greater and contour bulge or irregularity, and grade 3 refers to gross visible extension beyond the prostate. Grade 1, grade 2, and grade 3 had sensitivities of 24%, 36%, and 66%, respectively, for detection of EPE at final pathologic examination (5). These features were evaluated statistically to optimize their value in the grading system. Incorporating prostate-specific antigen level and Gleason score improved performance of the grading system, although there was no exact prescription regarding how this information is to be added to EPE grade. Thus, the grading system proposed by Mehralivand et al is currently an imaging-only system.

One striking aspect of the EPE grading system is how poor multiparametric MRI is at predicting EPE. Even in the case of gross visible extension, only 66% of cases proved to have EPE pathologically. Thus, EPE should always be diagnosed with some degree of humility at multiparametric MRI, and referring clinicians should tune their expectations accordingly. There are several explanations for the lack of predictive value. False-positive findings due to inflammation, desmoplastic reaction, and trauma-related changes from biopsy can lead to irregularities in the apparent margin of the tumor, thus simulating EPE. Meanwhile, at the other end of the spectrum, EPE is commonly microscopic and thus, below the detection threshold of multiparametric MRI (6). Moreover, contact length and

From the Molecular Imaging Program, National Cancer Institute, 10 Center Dr, Bldg 10, Room B3B69F, Bethesda, MD 20892. Received October 29, 2019; accepted October 30. **Address correspondence to** the author (e-mail: *pchoyke@nih.gov*).

See also the article by Reisæter et al in this issue. Conflicts of interest are listed at the end of this article.

bulging may be just that, containment of the tumor within the confines of the prostate. In short, if the proposed EPE grading system brings a systematic approach for EPE evaluation and reminds us that multiparametric MRI is not always highly predictive of EPE, then it will have accomplished a great deal. Such a system can potentially increase the awareness of the surgeons about the risk of EPE, can alter patient treatment, and will play an important role to convey such risks in reports in an organized fashion.

The current article compares a Likert system in use at the author's institution (7) to the EPE grading system of Mehralivand et al and finds it comparable in performance. To remind readers, the Likert system is a gestalt of various features summarized in a five-point scale: 1 = criterion not present, 2 = probably not present, 3 = uncertain if present, 4 = probably present, 5 = definitely present. While this system works as well as the EPE grading system of Mehralivand et al, it is understandably more difficult to teach, convey, and report on than a simpler three-grade system with stringently defined criteria and outcome measures. However, in this study of more than 300 patients, the results were supportive of the Mehralivand EPE grading system. Average sensitivity across both readers was 92.5% for EPE Mehralivand grade versus 77% for EPE Likert score.

Incorporating a standardized EPE reporting system such as the Mehralivand grade into PI-RADS would be helpful and further improve the uniformity of multiparametric MRI interpretations. If the grading system outcomes were placed as a footnote within reports, the grade would be readily associated with its corresponding quantitative risk of EPE.

There are several final points to be made about the comparison of the EPE grading system by Mehralivand et al and Likert scores for EPE. First, neither grading system corrects for reader variability. The criteria for both systems were equally subjective and thus, it is not surprising that readers will differ with each other to a similar extent. The κ value of 0.47 (fair) in the current study is approximately the same as that for PI-RADS itself (8). When comparing readings for each radiologist, there was a trend toward increased sensitivity at the cost of decreased specificity of Mehralivand EPE grade. It is arguable whether sensitivity or specificity is more important in this setting and further discussion on this point is warranted. Improvements in interreader variability will likely have to await the addition of machine learning methods that provide computational assessments of the risk of EPE in a standardized manner. Unfortunately, we are a long way from such systems.

Finally, the authors point out that many pathologists have moved to a more quantitative assessment of EPE, measuring the actual radial extension distance of the tumor from the prostate on histologic slides (3). Future iterations of the EPE grading system might correlate outcomes with radial excursion distance rather than simply the presence or absence of EPE.

In closing, as templated structured reporting becomes more common, it should also be applied to multiparametric MRI staging of prostate cancer. Mehralivand et al showed that EPE results could be improved by incorporating clinical risk features such as tumor grade, serum prostate-specific antigen level, and age into the risk assessment. Nevertheless, the Mehralivand EPE grading system, even as a stand-alone imaging-based system, could readily be incorporated into PI-RADS on its next iteration. It seems to perform as well as other systems but with the advantage of standardization. We encourage radiologists to utilize such repeatable grading systems, not only for prostate cancer, but for other oncologic imaging, to improve patient assessments.

Disclosures of Conflicts of Interest: P.L.C. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed no relevant relationships. Other relationships: author's institution issued patent on MRI transrectal US fusion biopsy; author's institution licensed to In Vivo; author receives royalties from the U.S. government.

References

1. Ball MW, Partin AW, Epstein JI. Extent of extraprostatic extension independently influences biochemical recurrence-free survival: evidence for further pT3 subclassification. Urology 2015;85(1):161–164.
2. Epstein JI, Carmichael MJ, Pizov G, Walsh PC. Influence of capsular penetration on progression following radical prostatectomy: a study of 196 cases with long-term followup. J Urol 1993;150(1):135–141.
3. Sung MT, Lin H, Koch MO, Davidson DD, Cheng L. Radial distance of extraprostatic extension measured by ocular micrometer is an independent predictor of prostate-specific antigen recurrence: A new proposal for the substaging of pT3a prostate cancer. Am J Surg Pathol 2007;31(2):311–318.
4. Wheeler TM, Dillioglugil O, Kattan MW, et al. Clinical and pathological significance of the level and extent of capsular invasion in clinical stage T1-2 prostate cancer. Hum Pathol 1998;29(8):856–862.
5. Mehralivand S, Shih JH, Harmon S, et al. A Grading System for the Assessment of Risk of Extraprostatic Extension of Prostate Cancer at Multiparametric MRI. Radiology 2019;290(3):709–719.
6. Costa DN, Passoni NM, Leyendecker JR, et al. Diagnostic Utility of a Likert Scale Versus Qualitative Descriptors and Length of Capsular Contact for Determining Extraprostatic Tumor Extension at Multiparametric Prostate MRI. AJR Am J Roentgenol 2018;210(5):1066–1072.
7. Reisæter LA, Halvorsen OJ, Beisland C, et al. Assessing extraprostatic extension with multiparametric MRI of the prostate: Mehralivand extraprostatic extension grade or extraprostatic extension Likert score? Radiol Imaging Cancer 2020;2(1):e190071.
8. Gaur S, Harmon S, Mehralivand S, et al. Prospective comparison of PI-RADS version 2 and qualitative in-house categorization system in detection of prostate cancer. J Magn Reson Imaging 2018;48(5):1326–1335.