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


Editorial Comment

Recent prospective studies, including randomized clinical trials, have provided compelling evidence in favor of the use of multi-parametric 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