Prostate Cancer: Can Multiparametric MR Imaging Help Identify Patients Who Are Candidates for Active Surveillance?[1]

Baris Turkbey, MD
Haresh Mani, MD
Omer Aras, MD
Jennifer Ho
Anthony Hoang, MD
Ardeshir R. Rastinehad, DO
Harsh Agarwal, PhD
Vijay Shah, PhD
Marcelino Bernardo, BS
Yuxi Pang, PhD
Dagane Daar, RT
Yolanda L. McKinney, RN
W. Marston Linehan, MD
Aradhana Kaushal, MD
Maria J. Merino, MD
Bradford J. Wood, MD
Peter A. Pinto, MD
Peter L. Choyke, MD

Purpose: To determine whether multiparametric magnetic resonance (MR) imaging can help identify patients with prostate cancer who would most appropriately be candidates for active surveillance (AS) according to current guidelines and to compare the results with those of conventional clinical assessment scoring systems, including the D'Amico, Epstein, and Cancer of the Prostate Risk Assessment (CAPRA) systems, on the basis of findings at prostatectomy.

Materials and Methods: This institutional review board–approved HIPAA-compliant retrospectively designed study included 133 patients (mean age, 59.3 years) with a mean prostate-specific antigen level of 6.73 ng/mL (median, 4.39 ng/mL) who underwent multiparametric MR imaging at 3.0 T before radical prostatectomy. Informed consent was obtained from all patients. Patients were then retrospectively classified as to whether they would have met AS eligibility criteria or were better served by surgery. AS eligibility criteria for prostatectomy specimens were a dominant tumor smaller than 0.5 mL without Gleason 4 or 5 patterns or extracapsular or seminal vesicle invasion. Conventional clinical assessment scores (the D'Amico, Epstein, and CAPRA scoring systems) were compared with multiparametric MR imaging findings for predicting AS candidates. The level of significance of difference between scoring systems was determined by using the χ^2 test for categoric variables with the level of significance set at $P < .05$.

Results: Among 133 patients, 14 were eligible for AS on the basis of prostatectomy results. The sensitivity, positive predictive value (PPV), and overall accuracy, respectively, were 93%, 25%, and 70% for the D'Amico system, 64%, 45%, and 88% for the Epstein criteria, and 93%, 20%, and 59% for the CAPRA scoring system for predicting AS candidates ($P < .005$ for all, χ^2 test), while multiparametric MR imaging had a sensitivity of 93%, a PPV of 57%, and an overall accuracy of 92% ($P < .005$).

Conclusion: Multiparametric MR imaging provides useful additional information to existing clinicopathologic scoring systems of prostate cancer and improves the assignment of treatment (eg, AS or active treatment).

Supplemental material: *http://radiology.rsna.org/lookup/suppl/doi:10.1148/radiol.13121325/-/DC1*

[1] From the Molecular Imaging Program (B.T., O.A., M.B., D.D., Y.L.M., P.L.C.), Laboratory of Pathology (H.M., M.J.M.), Radiation Oncology Branch (J.H., A.K.), and Urologic Oncology Branch (A.H., A.R.R., W.M.L., P.A.P.), National Cancer Institute, National Institutes of Health, 10 Center Dr, MSC 1182 Bldg 10, Room B3B69, Bethesda, MD 20892-1088; Philips Research North America, Briarcliff Manor, NY (H.A.); VirtualScopics, Rochester, NY (V.S.); Department of Imaging Physics, SAIC Frederick, National Cancer Institute-Frederick, Frederick, Md (M.B., D.D.); Philips Healthcare, Cleveland, Ohio (Y.P.); and Center for Interventional Oncology, National Cancer Institute and Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda, Md (B.J.W.). Received June 16, 2012; revision requested August 16; revision received December 6; accepted December 21; final version accepted December 28. **Address correspondence to** P.L.C. (e-mail: *pchoyke@mail.nih.gov*).

Prostate cancer is the most common cancer among men in the Western world, with the highest incidence and the third highest mortality rate among malignancies (1). In the United States, the number of estimated new cases and deaths per year are 241 740 and 28 170, respectively (2). Screening with prostate-specific antigen (PSA) levels has led to an increased incidence of prostate cancer, but these "screening" cancers are generally smaller and of lower grade and stage than clinically detected cancers, leading to fears of overtreatment (3,4). The aim of active surveillance (AS) is to avoid radical treatment unless disease progression occurs or the individual with prostate cancer decides to undergo treatment. AS has become an acceptable mode of treatment, but concern remains that the patient's tumor may actually be more aggressive than originally thought on the basis of increasing PSA levels, inconsistencies between different scoring systems, and concerns about undersampling during random prostate biopsies (5–7). Increasingly, such patients are initially offered AS, but they often switch to active treatment (AT) out of concerns that their disease is being underestimated (8–10). Currently, several criteria are in use that incorporate clinical-pathologic criteria such as serum PSA levels, PSA density and kinetics, digital rectal examination findings, and the number of cancer-positive cores with Gleason scores (together with the percentage of cores) at biopsy (11). Although promising results have been published by several groups, the accurate characterization of disease extent remains a source of concern when committing a patient to AS. Moreover, the situation is confounded by a considerable rate of misclassification and inconsistency when current risk assessment schemes are utilized (12–16). A major concern is that results of random 12-core biopsies do not accurately reflect the aggressiveness of the disease. Indeed, even extended biopsy protocols can miss cancers with unfavorable features, leading to improper selection for AS (17).

Multiparametric magnetic resonance (MR) imaging, including both anatomic and functional sequences, has been shown to be effective for the detection and local staging of prostate cancer (18); however, multiparametric MR imaging currently is not included in the decision-making algorithms or criteria for AS. The purpose of this study was to determine whether multiparametric MR imaging can help identify patients who would most appropriately be candidates for AS according to current guidelines and to compare the results with those of conventional clinical assessment scoring systems, including the D'Amico, Epstein, and Cancer of the Prostate Risk Assessment (CAPRA) systems, on the basis of findings at prostatectomy.

Advance in Knowledge

- On the basis of final histopathologic outcomes, incorporation of multiparametric MR imaging into the D'Amico, Epstein, or Cancer of the Prostate Risk Assessment scoring systems would have dramatically reduced (by 85%, 75%, and 91%, respectively) the number of misclassifications by these systems in assigning patients with prostate cancer to active treatment or active surveillance.

Materials and Methods

Study Design and Population

This retrospective single-institution study at National Cancer Institute, National Institutes of Health (Bethesda, Md) was approved by the local institutional review board and was compliant with the Health Insurance Portability and Accountability Act; informed consent was obtained from each patient.

Implication for Patient Care

- Multiparametric MR image evaluation can be used in determining candidates for active surveillance versus active treatment for prostate cancer.

One hundred forty-three patients (mean age, 59.4 years; median age, 59.0 years; range, 39–74 years) with a mean serum PSA level of 6.55 ng/mL (median, 4.7 ng/mL; range, 0.9–48.9 ng/mL) were enrolled in the study between January 2007 and August 2010. To be included in the study, patients had to have clinical-pathologic parameters available for the assessment of AS eligibility on the basis of the Epstein criteria (endorsed by the National Comprehensive Cancer Network), the CAPRA system, and the D'Amico criteria. In addition, the patient had to have had undergone a preoperative multiparametric MR imaging examination (with at least three of four of the following available sequences: triplane T2-weighted MR imaging, diffusion-weighted [DW] MR imaging, MR spectroscopy, and dynamic contrast material–enhanced MR imaging) at 3.0 T, followed by

Published online before print
10.1148/radiol.13121325 **Content codes:** GU OI

Radiology 2013; 268:144–152

Abbreviations:
AS = active surveillance
AT = active treatment
CAPRA = Cancer of the Prostate Risk Assessment
DW = diffusion weighted
PPV = positive predictive value
PSA = prostate-specific antigen
RARP = robotic-assisted radical prostatectomy

Author contributions:
Guarantors of integrity of entire study, B.T., H.A., P.L.C.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, B.T., H.M., O.A., A.H., A.R.R., H.A., M.J.M.; clinical studies, B.T., H.M., J.H., A.H., A.R.R., H.A., M.B., Y.P., D.D., Y.L.M., W.M.L., M.J.M., B.J.W., P.L.C.; experimental studies, B.T., H.M., A.R.R., H.A., V.S., M.B.; statistical analysis, B.T., O.A., H.A.; and manuscript editing, B.T., H.M., O.A., A.H., A.R.R., H.A., V.S., M.B., Y.P., Y.L.M., W.M.L., M.J.M., B.J.W., P.L.C.

Funding:
B.T., H.M., O.A., J.H., A.H., A.R.R., M.B., D.D., Y.L.M., W.M.L., A.K., M.J.M., B.J.W., P.A.P., and P.L.C. are employees of the National Institutes of Health. This research was supported by the National Institutes of Health (grants Z01 BC 011023 and Z01 BC 011081).

Conflicts of interest are listed at the end of this article.

Figure 1

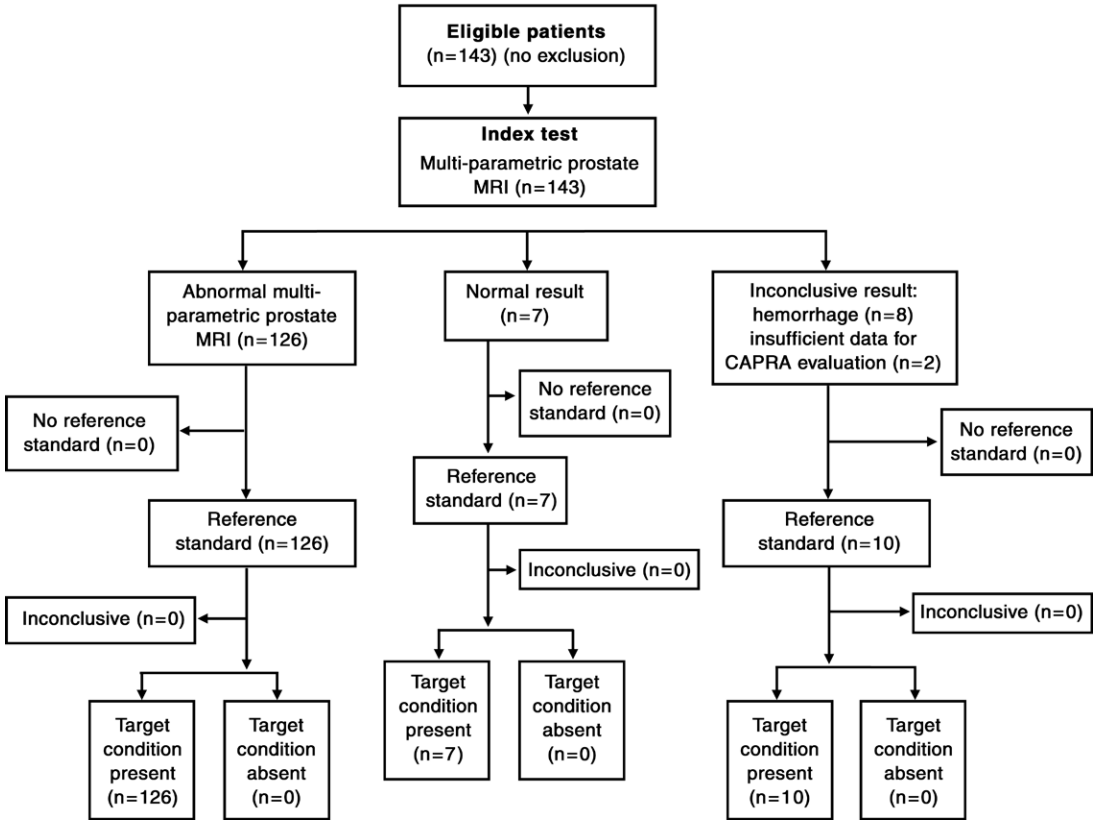

Figure 1: Flowchart of patient population.

robotic-assisted radical prostatectomy (RARP). Eight patients were excluded because their MR imaging studies could not be evaluated because of biopsy-related residual hemorrhage; an additional two patients were excluded owing to insufficient clinical-pathologic data for the CAPRA criteria.

The final patient population consisted of 133 patients (mean age, 59.3 years; median, 59.0 years; range, 39–74 years) with a mean PSA level of 6.73 ng/mL (median, 4.39 ng/mL; range, 0.9–48.9 ng/mL) (Fig 1). Clinical-pathologic features in the study population are presented in Table 1.

Multiparametric MR Imaging

All MR imaging examinations were performed by using a combination of an endorectal coil (BPX-30; Medrad, Pittsburgh, Pa) tuned to 127.8 MHz and a six- or 16-channel cardiac coil (SENSE; Philips Medical Systems, Best, the Netherlands) with a 3.0-T magnet (Achieva; Philips Medical Systems), without prior bowel preparation. The endorectal coil was inserted by using a semianesthetic gel (xylocaine, Lidocaine; AstraZeneca, Wilmington, Del) while the patient was in the left lateral decubitus position. The balloon surrounding the coil was distended with 3-mol/L perfluorocarbon (Fluorinert; 3M, St Paul, Minn) to a volume of approximately 50 mL to reduce susceptibility artifacts induced by air in the coil's balloon. The MR imaging protocol included triplanar T2-weighted turbo spin-echo imaging, DW MR imaging, three-dimensional MR spectroscopy, axial precontrast T1-weighted MR imaging, and axial three-dimensional fast-field-echo dynamic contrast-enhanced MR imaging. Axial dynamic contrast-enhanced images were obtained before, during, and after a single-dose injection of gadopentetate dimeglumine (Magnevist; Berlex, Wayne, NJ) administered at a dose of 0.1 mmol per kilogram of body weight through a peripheral vein at a rate of 3 mL/sec by using a mechanical injector (Spectris MR Injection System; Medrad). Sequence parameters were defined in previous studies (18,19) (Table 2). Among 133 patients in the final study population, 50 underwent multiparametric MR imaging with three pulse sequences (T2-weighted MR imaging, MR spectroscopy, and dynamic contrast-enhanced MR imaging) because DW MR imaging was not applied routinely at the time this cohort was imaged, while the remaining 83 patients underwent multiparametric MR imaging with four pulse sequences (T2-weighted MR imaging, DW MR imaging, MR spectroscopy, and dynamic contrast-enhanced MR imaging).

Table 2

Multiparametric MR Imaging Pulse Sequences

Sequence	Repetition Time (msec)/Echo Time (msec)	Field of View (mm)	Pixel Size (mm)	Matrix	Flip Angle(s) (degrees)	Section Thickness (mm)	Imaging Time
T2-weighted sagittal	2925/120	140	0.27 × 0.27	304 × 234	90/100	3	1 Min 51 sec
T2-weighted axial	8869/120	140	0.27 × 0.27	304 × 234	90/180	3	5 Min 37 sec
T2-weighted coronal	2632/120	140	0.27 × 0.27	304 × 234	90/100	3	1 Min 40 sec
DW imaging	3709/52	140	1.02 × 1.02	112 × 108	90/180	2.73/0.27	4 Min 46 sec
3D MR spectroscopy	980/100	72	6 × 6	10 × 10	90/180	6	12 Min 51 sec
T1-weighted axial GRE	3.7/2.2	262	1.02 × 1.02	256 × 186	2	3	0.22 Sec
T1-weighted axial dynamic contrast enhanced	3.7/2.2	262	1.02 × 1.02	256 × 186	8.5	3	5 Min 12 sec
Axial THRIVE	5.3/2.6	440	1.5 × 1.5	280 × 199	10	5	0.57 Sec

Note.—DW imaging was performed in 83 patients (at DW imaging, five evenly spaced *b* values between 0 and 750 sec/mm^2 were used). GRE = gradient echo, 3D = three-dimensional, THRIVE = T1 high resolution isotropic volume examination.

Table 1

Clinical-pathologic Features in the Study Population (*n* = 133)

Feature	Datum
Age (y)	
Mean	59.3
Median	59.0
Range	39–74
Serum PSA level (ng/mL)	
Mean	6.73
Median	4.39
Range	0.9–48.9
Clinical tumor stage*	
T1c	96 (72)
T2a	29 (22)
T2b	8 (6)
Tumor Gleason score at biopsy*	
3 + 3	54 (40.6)
3 + 4	53 (39.8)
4 + 3	9 (6.8)
4 + 4	9 (6.8)
4 + 5	4 (3.0)
5 + 4	1 (0.8)
5 + 5	1 (0.8)
Not available	2 (1.5)

* Data are numbers of patients, with percentages in parentheses. Percentages may not add up to 100% because of rounding.

Histopathologic Examination

All patients underwent RARP under the direction of a single surgeon (P.A.P., with more than 10 years of experience in prostatectomy). The mean interval between MR imaging and RARP was 60 days (range, 3–180 days; median, 48 days). In the first set of 60 patients, the specimens were sectioned manually from apex to base at 4-mm intervals after RARP. In the subsequent 73 patients, a customized MR imaging–based specimen mold system was used to slice the prostatectomy specimens. In that system, the specimen was fixed in formalin for 2–24 hours at room temperature and then was placed in the customized three-dimensional mold and sliced in axial 6-mm sections. In both techniques, each slice was sequentially annotated by slice number and was then kept in fixative for a further period of 24–48 hours, followed by paraffin embedding, cutting, and mounting on a whole-mount glass slide. Whole-mount sections were 5 μm thick.

Data Analysis

The dominant tumor was outlined on each prostatectomy specimen by two experienced genitourinary pathologists (H.M. and M.J.M., with 10 and 25 years of experience, respectively) who were blinded to the MR imaging data. The criteria for the dominant lesion were that it was the largest and the most aggressive focus (on the basis of Gleason score) for that patient (20). The volume of each outlined tumor was measured by using the ellipsoid formula (length times width times height times 0.52) in cubic centimeters (eg, a tumor with greatest axial dimensions of 0.8 × 1.0 cm outlined on a single 3-mm slice was measured as 0.8 × 1.0 × 0.3 × 0.52 = 0.124 cm^3). Extracapsular extension and seminal vesicle invasion were assessed for each specimen.

For MR image analysis, two experienced genitourinary radiologists (B.T. and P.L.C., with 5 and 11 years of experience in prostate MR imaging, respectively) evaluated T2-weighted MR images, apparent diffusion coefficient maps from DW MR imaging, MR spectroscopic images, and dynamic contrast-enhanced MR images in consensus. The reviewers were blinded to the clinical-pathologic findings (PSA levels, clinical stage findings, and biopsy findings) and to the histopathologic results. The dominant tumor for each patient was determined on the basis of its size. An imaging score was assigned to each lesion on the basis of its features on images obtained with different pulse sequences at MR imaging, yielding low, moderate, and high suspicion levels (Table 3). For multiparametric MR imaging analysis, on T2-weighted MR images and apparent diffusion coefficient maps from DW MR imaging, the criterion for a "visible" lesion was a well circumscribed, round-ellipsoid, low-signal-intensity region within the prostate gland (18,19). The analysis of three-dimensional MR spectroscopy images evaluated choline (Cho)/citrate (Cit) ratios within voxels in the biopsy core sites. The mean healthy Cho/Cit

ratio was defined as 0.13 ± 0.081 on the basis of previously reported results in 433 healthy voxels from peripheral zone regions with negative biopsy results in 44 additional patients who were referred for prostate MR imaging and who had histologic confirmation. Voxels were considered abnormal when the Cho/Cit ratio was 3 or more standard deviations higher than the mean healthy Cho/Cit ratio (0.373) (18,19). Dynamic contrast-enhanced MR images were evaluated by direct visual interpretation of raw dynamic contrast-enhanced T1-weighted images, where the diagnostic criterion for prostate cancer was a focus of early and intense enhancement with rapid washout compared with the background (18,19). One of the readers (B.T.) manually segmented the dominant tumor at a picture archiving and communication system (PACS) workstation (Carestream Health, Rochester, NY). The tumor volume was automatically obtained from the PACS software after lesions were contoured on MR images. For segmenting the tumors, T2-weighted MR images, apparent diffusion coefficient maps from DW MR imaging, and dynamic contrast-enhanced MR images were used to determine tumor boundaries in combination, although the final regions of interest were superimposed on T2-weighted MR images.

The criteria used for AS eligibility regarding prostatectomy specimens were those previously defined by Duffield et al (21) and included having a dominant tumor smaller than 0.5 mL without a Gleason 4 or 5 pattern or extracapsular or seminal vesicle invasion. Criteria for AS at multiparametric MR imaging were having a dominant tumor smaller than 0.5 mL without extracapsular extension or seminal vesicle invasion and a low imaging score (positive findings only on T2-weighted MR images and/or DW MR images or only on MR spectroscopy images or only on dynamic contrast-enhanced MR images) (Table 4). Any patient whose tumor did not meet all of these criteria was considered to be ineligible for AS.

For each patient, clinical-pathologic data (serum PSA level, serum PSA density, clinical stage based on digital rectal examination findings, number of positive biopsy cores, percentage of positive biopsy cores, percentage of involvement in positive biopsy cores) were used to assess eligibility for AS versus AT for the D'Amico, Epstein, and CAPRA criteria (Table 4). The purpose of each clinical scoring system is to assign risk to the patient, thereby helping determine if the patient may be more appropriately treated with AS or with AT. The prediction of low risk suitable for AS was assigned first without MR imaging results and was then compared with MR imaging results. These results were compared with the histopathologic determination of risk and suitability for AS.

Table 3

Multiparametric MR Image Evaluation Score Chart

T2-weighted MR Imaging	Apparent Diffusion Coefficient Maps from DW MR Imaging	MR Spectroscopy	Dynamic Contrast-enhanced MR Imaging	Multiparametric MR Imaging Suspicion Level
Negative	Negative	Negative	Negative	Negative
Positive	Negative	Negative	Negative	Low
Positive	Positive	Negative	Negative	Low
Negative	Positive	Negative	Negative	Low
Negative	Negative	Positive	Negative	Low
Negative	Negative	Negative	Positive	Low
Positive	Negative	Positive	Negative	Moderate
Positive	Negative	Negative	Positive	Moderate
Negative	Positive	Positive	Negative	Moderate
Negative	Positive	Negative	Positive	Moderate
Positive	Positive	Positive	Negative	Moderate
Positive	Positive	Negative	Positive	Moderate
Negative	Negative	Positive	Positive	Moderate
Positive	Positive	Positive	Positive	High

Note.—The criteria for positive and negative findings on each kind of MR image are described in the Data Analysis section of the text.

Table 4

AS Criteria for Clinical-pathologic, Multiparametric MR Imaging, and Whole-Mount Histopathologic Examination Approaches

Approach	Criteria for AS Eligibility
D'Amico scoring system	Serum PSA level ≤ 10 ng/mL, clinical stage T2a or lower, no Gleason pattern 4 or 5 at biopsy
Epstein criteria	Serum PSA density < 0.15 ng/mL, clinical stage T1c, no Gleason pattern 4 or 5 at biopsy, fewer than three positive cores at biopsy, <50% cancer involvement per core
CAPRA scoring system	Patient age, serum PSA level at diagnosis, primary and secondary Gleason patterns at biopsy, clinical (T) stage, and percentage of cancer involvement at biopsy*
Multiparametric MR imaging scoring system	Dominant tumor volume < 0.5 mL, low suspicion score at multiparametric MR imaging, no extracapsular extension, no seminal vesicle invasion
Whole-mount histopathologic examination system	Dominant tumor volume < 0.5 mL, no Gleason 4 or 5 pattern, no extracapsular extension, no seminal vesicle invasion

* A CAPRA score of 0–2 indicates low risk *(http://urology.ucsf.edu/patientGuides/uroOncPt_Assess.html#capra).*

Figure 2

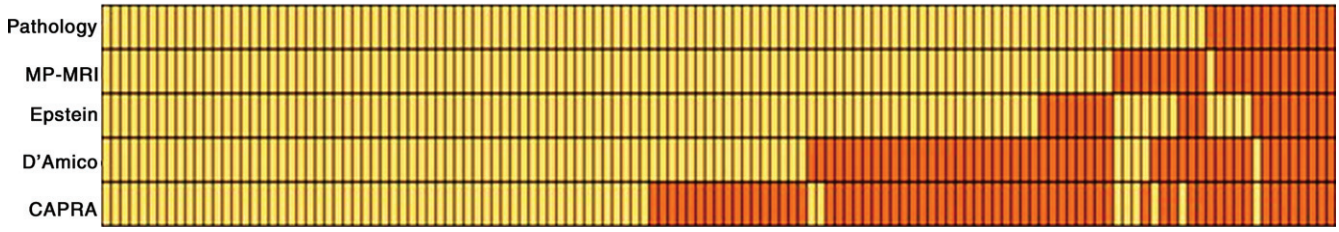

Figure 2: Array shows results of AT (yellow) and AS (orange) decision making by using pathologic examination results (at the top row as the reference standard), multiparametric *(MP)* MR imaging, the Epstein criteria, the D'Amico system, and the CAPRA approach. Each patient is represented by a column of the array.

Statistical Analysis

A bivariate analysis (χ^2 test or Fisher exact test) was used to determine whether there was a correlation between the MR imaging scoring system and the D'Amico, Epstein, and CAPRA scoring systems. The level of statistical significance in this study was set at $P <$ .05. All analyses were performed with statistical software (SPSS, version 16.0; SPSS, Chicago, Ill).

Results

At histopathologic examination, a total of 133 dominant lesions were identified in 133 patients (16 Gleason 3 + 3 lesions, 72 Gleason 3 + 4 lesions, four Gleason 3 + 5 lesions, six Gleason 4 + 3 lesions, 17 Gleason 4 + 4 lesions, 16 Gleason 4 + 5 lesions, and two Gleason 5 + 4 lesions). Extracapsular extension and seminal vesicle invasion were present in 46 and six patients, respectively.

Multiparametric MR imaging depicted 126 of 133 dominant lesions. In seven patients, multiparametric MR imaging did not depict any lesion, and histopathologic evaluation in these patients demonstrated Gleason 3 + 3 tumor in five patients (tumor volume range, 0.004–0.42 mL), Gleason 4 + 3 tumor in one patient (tumor volume, 0.13 mL), and Gleason 4 + 5 tumor in one patient (tumor volume, 0.98 mL). The sensitivity, specificity, positive predictive value (PPV), and negative predictive value of multiparametric MR imaging for extracapsular extension were 76% (35 of 46), 95% (83 of 87), 90% (35 of 39), and 88% (83 of 94), respectively. The sensitivity, specificity, PPV, and negative predictive value of multiparametric MR imaging for seminal vesicle invasion were 67% (four of six), 100% (127 of 127), 100% (four of four), and 98% (127 of 129), respectively.

On the basis of histopathologic criteria, 14 of 133 patients were found to meet established criteria for AS. The sensitivity, PPV, and overall accuracy of the D'Amico scoring system for predicting AS candidates were 93% (13 of 14), 25% (13 of 52), and 70% (93 of 133) ($P <$ .0001), respectively (Table E1 [online]). The sensitivity, PPV, and overall accuracy of the Epstein criteria for predicting AS candidates were 64% (nine of 14), 45% (nine of 20), and 88% (117 of 133) ($P <$ .0001), respectively (Table E2 [online]). The sensitivity, PPV, and overall accuracy of the CAPRA scoring system for predicting AS candidates were 93% (13 of 14), 20% (13 of 66), and 59% (79 of 133) (P = .001), respectively (Table E3 [online]). The sensitivity, PPV, and overall accuracy of multiparametric MR imaging for predicting AS candidates were 93% (13 of 14), 57% (13 of 23), and 92% (122 of 133) ($P <$ .0001), respectively (Table E4 [online]). Multiparametric MR image evaluation led to the misclassification of 11 of 133 patients (n = 1 eligible for AS; n = 10 eligible for AT). The PPVs of the D'Amico scoring system, the Epstein criteria, the CAPRA score, and multiparametric MR imaging for predicting AT candidates were 99% (80 of 81), 96% (108 of 113), 99% (66 of 67), and 99% (109 of 110), respectively (Fig 2).

With the D'Amico scoring system, 40 of 133 patients were misclassified (n = 1 eligible for AS; n = 39 eligible for AT). Among the 39 misclassified AT candidates, eight were classified as AS candidates because their disease had a clinical stage of T2a, according to the D'Amico scoring system. The remainder of misclassifications of the AT candidates were secondary to undersampling or underscoring at biopsy. One AS candidate who was misclassified as an AT candidate had a serum PSA level of 14.6 ng/mL; however, the prostate volume in this patient at MR imaging was 153 mL, resulting in a PSA density of 0.11 ng/mL (<0.15 ng/mL). The elevated PSA in this patient could therefore be explained by prostatic hyperplasia, and multiparametric MR image evaluation resulted in correctly classifying him as an AS candidate. Incorporation of multiparametric MR image evaluation into the D'Amico scoring system would have corrected the classification of 34 (85%) of these 40 misclassified patients (n = 1 eligible for AS, n = 33 eligible for AT) (Table E5 [online]).

With the Epstein criteria, 16 of 133 patients were misclassified (n = 5 eligible for AS; n = 11 eligible for AT). The misclassification of all misclassified AT candidates was due to undersampling or underscoring at biopsy, whereas the misclassification of AS candidates for AT was due to high serum PSA density and overstaging in two and three patients, respectively. Incorporation of multiparametric MR image evaluation into the Epstein criteria corrected misclassification in 12 (75%) of these 16 patients (n = 4 eligible for AS, n = 8 eligible for AT) (Table E6 [online]).

With the CAPRA scoring system, 54 of 133 patients were misclassified

(n = 1 eligible for AS; n = 53 eligible for AT). Misclassification in all 53 misclassified AT candidates was due to undersampling or underscoring at biopsy, whereas one AS candidate misclassified for AT had an enlarged prostate gland measuring 153 mL with a serum PSA level of 14.6 ng/mL. Incorporation of multiparametric MR image evaluation into the CAPRA scoring system corrected misclassification in 49 (91%) of these 54 patients (n = 1 eligible for AS, n = 48 eligible for AT) (Table E7 [online]).

Discussion

Accurate stratification of patients into AS or AT for prostate cancer management is critical to successful implementation of an AS program and to reduce the morbidity associated with whole-gland therapies. Most conversions of AS to AT occur 1–2 years after the original assignment and are usually the result of misclassification due to undersampled and/or underscored clinicopathologic results (21). Our results suggest that multiparametric MR image evaluation can improve the identification of AS-eligible patients with prostate cancer compared with commonly used clinical-pathologic criteria. Moreover, when MR imaging was used in conjunction with clinical-pathologic criteria–based systems, it greatly improved the sensitivity and accuracy of each system.

Three clinical-pathologic criteria systems are in wide use. Among these, the D'Amico scoring system led to misclassification of 30% of our cohort, and this mainly included underestimation of the disease. Suardi et al (12) studied the D'Amico scoring system in a cohort of 2345 patients and reported a misclassification rate of 14%. They pointed out the importance of more selective criteria for AS. Rice et al (22) conducted a trial in 770 patients, 324 of whom were undergoing AS on the basis of the D'Amico criteria. Patients were followed up for up to 6.4 years. The poorest overall survival was for patients undergoing AS without any treatment. Recently, O'Brien et al (23) reported less favorable outcomes following prostatectomy after a delay of 6 months in patients who met D'Amico criteria for AS. Our work indicates that multiparametric MR image evaluation can assist the D'Amico scoring system in improving patient stratification for AS versus AT, as incorporation of multiparametric MR image evaluation was able to correct 87% of misclassifications.

The Epstein criteria, endorsed by the National Comprehensive Cancer Network guidelines, performed better in our study than the D'Amico and CAPRA scoring systems, with a misclassification rate of only 12%. This was most likely because of the strict nature of the Epstein criteria for AS compared with the other two scoring systems. Suardi et al (24) tested the ability of the Epstein criteria to help identify patients with low-risk prostate cancer suitable for AS and to correctly exclude unfavorable histopathologic features in a cohort of 874 patients. They reported ultimate unfavorable histopathologic findings in 3.3%–7.1% of patients who would have been eligible for AS on the basis of Epstein criteria. Lee et al (13) determined the performance of the Epstein criteria for predicting pathologic end points in men with early-stage prostate cancer treated with surgery. They reported that the Epstein criteria were able to help characterize insignificant disease in 34% of patients and concluded that the Epstein criteria can predict a high likelihood of organ-confined disease but not biologically indolent disease. Hekal et al (14) validated the Epstein criteria in a cohort of 35 patients and identified a Gleason score of greater than 6 and tumor Gleason score upgrading in 46% and 40% of patients, respectively, at final histopathologic examination. Recently, Tosoian et al (10) reported follow-up results in patients undergoing AS on the basis of the Epstein criteria. The median follow-up was 6.5 years, and the percentages of men remaining free of intervention after 2, 5, and 10 years of follow-up were 81%, 59%, and 41%, respectively. They emphasized the importance of limiting AS to patients with very low risk to reduce the adverse outcome frequency. Although the Epstein criteria achieved an overall accuracy of 88% in stratifying AS versus AT in our prostate cancer cohort, the criteria are inherently limited because of the undersampling of random biopsies, while incorporation of multiparametric MR imaging into the Epstein criteria improved the ability to stratify patients.

The CAPRA scoring system had a misclassification rate of 41%, which was mostly due to underestimation of disease. Behbahani et al (15) analyzed pathologic results in patients eligible for AS (on the basis of the CAPRA score) after radical prostatectomy in 125 patients and reported pathologic T2c cancer in 25.6% of the patients, whereas the Gleason score upgrade was 34.4% in that cohort. In our cohort, the CAPRA score system had a sensitivity of 55% for stratification of patients to AT; however, the incorporation of multiparametric MR imaging into the CAPRA system improved the sensitivity to 92% for AT stratification.

There are a limited number of studies incorporating MR imaging into clinical-pathologic decision algorithms. Guzzo et al (25) evaluated the ability of T2-weighted MR imaging findings to help predict adverse pathologic features in patients qualifying for AS on the basis of Epstein criteria and concluded that tumor identification at T2-weighted MR imaging was not predictive of adverse pathologic features in patients undergoing AS. However, this study did not include multiparametric MR imaging. Recently, Shukla-Dave et al (16) reported results of a newly designed nomogram that incorporates T2-weighted MR imaging and MR spectroscopy findings in 181 patients and concluded that the model nomogram improved the predictive accuracy for clinically unimportant prostate cancer, with areas under the curve that increased from 0.56 to 0.77 ($P <$ .001). Although the exact multiparametric MR imaging techniques and analytic methods used in our study are different from these published results, our findings also support the incorporation of multiparametric MR imaging findings into clinical-pathologic nomograms.

In this retrospective analysis, we found that MR imaging failed to depict

two lesions that at histopathologic examination proved to be dominant tumors of 0.13 mL (Gleason 4 + 3) and 0.98 mL (Gleason 4 + 5). In another patient, multiparametric MR imaging substantially underestimated the tumor volume as 0.27 mL, versus 4.6 mL at histopathologic examination. In seven patients, multiparametric MR imaging led to correct estimation of the tumor volumes; however, the tumors were greater than Gleason grade 6, despite their low volumes (<0.5 mL). It should also be noted that increasingly, Gleason 3 + 4 low-volume prostate cancers are considered for AS. van den Bergh et al (9) reported results of AS in 50 men with Gleason 7 prostate cancer. The 6-year treatment-free survival rate was 59%, with most patients switching to AT on the basis of their PSA levels. They proposed AS for patients with Gleason 3 + 4 prostate cancer, especially if they had comorbidities and/or short life expectancies. In our study, MR imaging led to the misclassification of only one AS candidate, when MR imaging findings overestimated the tumor volume as 0.9 mL, versus 0.4 mL at histopathologic examination.

Our study had several limitations. First, it was a relatively small retrospective cohort study conducted at a single institution; specifically, it should be noted that the number of patients who were eligible for AS was small. This indicates that most patients were appropriately treated, at least according to current guidelines. This is consistent with the experience of other groups as well (26). Clearly, our findings will need to be validated at other centers. Furthermore, in the initial 50 patients in the cohort, DW MR imaging was not performed, because that sequence was not applied routinely at the time this cohort was imaged. Our subsequent experience with DW MR imaging has shown that T2-weighted MR imaging and DW MR imaging findings quite often confirm each other (18). Another limitation was that the interpretation of the multiparametric MR images was subjective and binary (positive or negative). A standardized scoring system is currently being developed, but this assessment system is not yet validated for clinical use (27). Moreover, our study evaluated the consensus assessment of multiparametric MR images with two expert reviewers, which reflects the real-life scenario for decision making regarding AT or AS with multiparametric MR imaging. We did not evaluated interreader variability, but future work will assess individual reader variability with a larger number of cases and readers. Finally, we compared multiparametric MR image evaluation only with the most commonly used clinical-pathologic scoring systems (D'Amico, Epstein, and CAPRA); however, further work is needed to compare our findings with other nomograms.

In conclusion, AS is an alternative treatment option in patients with indolent prostate cancer, but current nomograms based on clinical-pathologic criteria often incorrectly stratify these patients to AS or AT. On the basis of this review of surgical outcomes, multiparametric MR imaging proved to be superior to the D'Amico, Epstein, and CAPRA scoring systems in correctly classifying patients for AS versus AT and greatly improved the accuracy of all of the scoring systems when it was combined with them. Multiparametric MR image evaluation, incorporating an MR imaging scoring system with dominant tumor volume measurement, could be helpful in stratifying patients with prostate cancer for AS or AT in conjunction with existing guidelines.

Disclosures of Conflicts of Interest: B.T. No relevant conflicts of interest to disclose. **H.M.** No relevant conflicts of interest to disclose. **O.A.** No relevant conflicts of interest to disclose. **J.H.** No relevant conflicts of interest to disclose. **A.H.** No relevant conflicts of interest to disclose. **A.R.R.** No relevant conflicts of interest to disclose. **H.A.** Financial activities related to the present article: is an employee of Philips Electronics. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **V.S.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: is an employee of VirtualScopics. Other relationships: none to disclose. **M.B.** No relevant conflicts of interest to disclose. **Y.P.** Financial activities related to the present article: is an employee of Philips Healthcare. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **D.D.** No relevant conflicts of interest to disclose. **Y.L.M.** No relevant conflicts of interest to disclose. **W.M.L.** No relevant conflicts of interest to disclose. **A.K.** No relevant conflicts of interest to disclose. **M.J.M.** No relevant conflicts of interest to disclose. **B.J.W.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: none to disclose. Other relationships: the National Institutes of Health and Philips Healthcare have a Cooperative Research and Development Agreement. **P.A.P.** No relevant conflicts of interest to disclose. **P.L.C.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: institution has research agreements with Siemens, GE Healthcare, and Philips. Other relationships: none to disclose.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61(2):69–90.
2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin 2012;62(1):10–29.
3. Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. Nat Rev Cancer 2008;8(4):268–278.
4. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 2009;360(13):1320–1328.
5. Cooperberg MR, Lubeck DP, Meng MV, Mehta SS, Carroll PR. The changing face of low-risk prostate cancer: trends in clinical presentation and primary management. J Clin Oncol 2004;22(11):2141–2149.
6. Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst 2002;94(13):981–990.
7. Seiler D, Randazzo M, Klotz L, et al. Pathological stage distribution in patients treated with radical prostatectomy reflecting the need for protocol-based active surveillance: results from a contemporary European patient cohort. BJU Int 2012;110(2):195–200.
8. Carter HB, Kettermann A, Warlick C, et al. Expectant management of prostate cancer with curative intent: an update of the Johns Hopkins experience. J Urol 2007;178(6):2359–2364; discussion 2364–2365.
9. van den Bergh RC, Roemeling S, Roobol MJ, et al. Outcomes of men with screen-detected prostate cancer eligible for active surveillance who were managed expectantly. Eur Urol 2009;55(1):1–8.
10. Tosoian JJ, Trock BJ, Landis P, et al. Active surveillance program for prostate cancer:

an update of the Johns Hopkins experience. J Clin Oncol 2011;29(16):2185–2190.

11. Bastian PJ, Carter BH, Bjartell A, et al. Insignificant prostate cancer and active surveillance: from definition to clinical implications. Eur Urol 2009;55(6):1321–1330.
12. Suardi N, Capitanio U, Chun FK, et al. Currently used criteria for active surveillance in men with low-risk prostate cancer: an analysis of pathologic features. Cancer 2008;113(8):2068–2072.
13. Lee MC, Dong F, Stephenson AJ, Jones JS, Magi-Galluzzi C, Klein EA. The Epstein criteria predict for organ-confined but not insignificant disease and a high likelihood of cure at radical prostatectomy. Eur Urol 2010;58(1):90–95.
14. Hekal IA, El-Tabey NA, Nabeeh MA, et al. Validation of Epstein criteria of insignificant prostate cancer in Middle East patients. Int Urol Nephrol 2010;42(3):667–671.
15. Behbahani TE, Ellinger J, Caratozzolo DG, Müller SC. Pathological outcomes of men eligible for active surveillance after undergoing radical prostatectomy: are results predictable? Clin Genitourin Cancer 2012;10(1):32–36.
16. Shukla-Dave A, Hricak H, Akin O, et al. Preoperative nomograms incorporating magnetic resonance imaging and spectroscopy for prediction of insignificant prostate cancer. BJU Int 2012;109(9):1315–1322.
17. Ploussard G, Salomon L, Xylinas E, et al. Pathological findings and prostate specific antigen outcomes after radical prostatectomy in men eligible for active surveillance: does the risk of misclassification vary according to biopsy criteria? J Urol 2010;183(2):539–544.
18. Turkbey B, Mani H, Shah V, et al. Multiparametric 3T prostate magnetic resonance imaging to detect cancer: histopathological correlation using prostatectomy specimens processed in customized magnetic resonance imaging based molds. J Urol 2011;186(5):1818–1824.
19. Turkbey B, Pinto PA, Mani H, et al. Prostate cancer: value of multiparametric MR imaging at 3 T for detection—histopathologic correlation. Radiology 2010;255(1):89–99.
20. Stamey TA, McNeal JE, Yemoto CM, Sigal BM, Johnstone IM. Biological determinants of cancer progression in men with prostate cancer. JAMA 1999;281(15):1395–1400.
21. Duffield AS, Lee TK, Miyamoto H, Carter HB, Epstein JI. Radical prostatectomy findings in patients in whom active surveillance of prostate cancer fails. J Urol 2009;182(5):2274–2278.
22. Rice KR, Colombo ML, Wingate J, et al. Low risk prostate cancer in men ≥ 70 years old: to treat or not to treat. Urol Oncol 2011 Aug 25. [Epub ahead of print]
23. O'Brien D, Loeb S, Carvalhal GF, et al. Delay of surgery in men with low risk prostate cancer. J Urol 2011;185(6):2143–2147.
24. Suardi N, Briganti A, Gallina A, et al. Testing the most stringent criteria for selection of candidates for active surveillance in patients with low-risk prostate cancer. BJU Int 2010;105(11):1548–1552.
25. Guzzo TJ, Resnick MJ, Canter DJ, et al. Endorectal T2-weighted MRI does not differentiate between favorable and adverse pathologic features in men with prostate cancer who would qualify for active surveillance. Urol Oncol 2012;30(3):301–305.
26. Silberstein JL, Vickers AJ, Power NE, et al. Reverse stage shift at a tertiary care center: escalating risk in men undergoing radical prostatectomy. Cancer 2011;117(21):4855–4860.
27. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. Eur Radiol 2012;22(4):746–757.