Likert Score 3 Prostate Lesions: Association Between Whole-Lesion ADC Metrics and Pathologic Findings at MRI/Ultrasound Fusion Targeted Biopsy

Andrew B. Rosenkrantz, MD,[1]* Xiaosong Meng, MD, PhD,[2] Justin M. Ream, MD,[1] James S. Babb, PhD,[1] Fang-Ming Deng, MD, PhD,[3] Henry Rusinek, PhD,[1] William C. Huang, MD,[2] Herbert Lepor, MD,[2] and Samir S. Taneja, MD[2]

Background: To assess associations between whole-lesion apparent diffusion coefficient (ADC) metrics and pathologic findings of Likert score 3 prostate lesions at MRI/ultrasound fusion targeted biopsy.
Methods: This retrospective Institutional Review Board-approved study received a waiver of consent. We identified patients receiving a highest lesion score of 3 on 3 Tesla multiparametric MRI reviewed by a single experienced radiologist using a 5-point Likert scale and who underwent fusion biopsy. A total of 188 score 3 lesions in 158 patients were included. Three-dimensional volumes-of-interest encompassing each lesion were traced on ADC maps. Logistic regression was used to predict biopsy results based on whole-lesion ADC metrics and patient biopsy history. Biopsy yield was compared between metrics.
Results: By lesion, targeted biopsy identified tumor in 22.3% and Gleason score (GS) >6 tumor in 8.5%, although results varied by biopsy history: biopsy-naïve (n = 80), 20.0%/8.8%; prior negative biopsy (n = 53), 9.4%/1.9%; prior positive biopsy (n = 55): 40.0%/14.5%. Biopsy history, whole-lesion mean ADC, whole-lesion $ADC_{10–25}$, and whole-lesion $ADC_{25–50}$ were each significantly associated with tumor or GS >6 tumor at fusion biopsy ($P \leq 0.047$). In men without prior negative prostate biopsy, whole-lesion $ADC_{25–50} \leq 1.04*10^{-3}$ mm^2/s achieved 90.0% sensitivity and 50.0% specificity for GS >6 tumor, which was significantly higher ($P < 0.001$) than specificity of PSA (17.5%) at identical sensitivity.
Conclusion: For score 3 lesions in patients without prior negative biopsy, whole-lesion ADC metrics help detect GS >6 cancer while avoiding negative biopsies. However, deferral of fusion biopsy may be considered for score 3 lesions in patients with prior negative biopsy (without applying whole-lesion ADC metrics) given exceedingly low ($\sim$ 2%) frequency of GS >6 tumor in this group.

J. MAGN. RESON. IMAGING 2016;43:325–332.

Prostate MRI has become a commonly applied test for detection and localization of clinically important tumors within the gland.[1] Increasing clinical adoption of prostate MRI has facilitated the emergence of MRI-targeted biopsy, whether using a direct in-bore approach[2] or MRI/ultrasound fusion guidance.[3] Recent studies support MRI-targeted biopsy in both patients with a prior negative biopsy[4,5] as well as in patients on active surveillance for biopsy-proven tumor,[2,6] and it is anticipated that the application of MRI-targeted prostate biopsy in clinical practice will continue to grow.

The varying level of concern for significant cancer among lesions encountered on prostate MRI creates a challenge in the clinical integration of MRI-targeted biopsy.[7] Some lesions are highly suspicious for aggressive tumor; others are overwhelmingly likely to be inconsequential. Thus, not all identified lesions warrant targeted biopsy. Given this variability, reporting schemes have been investigated to standardize

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.24983

Received Apr 24, 2015, Accepted for publication Jun 4, 2015.

*Address reprint requests to: A.B.R., Department of Radiology, Center for Biomedical Imaging, NYU School of Medicine, NYU Langone Medical Center, 660 First Avenue, 3rd Floor, New York, NY 10016. E-mail: andrew.rosenkrantz@nyumc.org

From the [1]Department of Radiology, NYU School of Medicine, NYU Langone Medical Center, New York, New York, USA; [2]Department of Urology, Division of Urologic Oncology, NYU School of Medicine, NYU Langone Medical Center, New York, New York, USA; and [3]Department of Pathology, NYU School of Medicine, NYU Langone Medical Center, New York, New York, USA

interpretations and communicate the radiologist's level of concern in a reproducible manner, which in turn may help guide biopsy selection.[8,9] The most frequently used approaches are the Likert and PI-RADS schemes, both of which in their most recent descriptions use a 1–5 scale for risk stratification.[10–12] Whereas PI-RADS uses fixed criteria for the assigned score, the Likert scheme uses the radiologist's overall impression. Regardless of the approach, uniform reporting with a 1–5 scale provides a consistent and straightforward means of conveying the findings as a basis for clinical decisions.

Past studies have investigated outcomes for prostate MRI examinations scored using 1–5 scales. Consistently, a score of 4 or 5 has been found to be highly sensitive for clinically significant tumor,[13,14] whereas a score of 1 or 2 has been found to have a high negative predictive value and to generally represent benign tissue or insignificant tumor.[2,14,15] Therefore, it has been suggested to routinely perform targeted biopsy of score 4 or 5 lesions, while considering deferring biopsy for low-probability lesions.[2]

As prior reports have largely drawn conclusions regarding management of score 1/2 or score 4/5 lesions, and not clearly committed to a position regarding biopsy of score 3 lesions, the proper management of such lesions is unclear. A score of 3 indicates an intermediate suspicion level, and data regarding outcomes of such lesions are variable.[12,16,17] Thus, strategies to further improve risk stratification among score 3 lesions and guide decisions regarding potential biopsy of such lesions would be of great clinical value.

Apparent diffusion coefficient (ADC) values of prostate lesions are associated with patient's risk profile, showing associations with Gleason score (GS),[18] posttreatment recurrences,[19] and outcomes on active surveillance.[20] Additional recent studies show further prognostic value of more sophisticated ADC metrics derived from whole-lesion histogram assessment.[21–23] Such metrics may also be useful for assisting management decisions in score 3 lesions. Thus, in this study, our aim was to assess associations between whole-lesion ADC metrics and pathologic findings of Likert score 3 prostate lesions at MRI/ultrasound fusion targeted biopsy.

Materials and Methods

Patients

This retrospective study was Health Insurance Portability and Accountability Act compliant and was approved by our institutional review board, which waived the requirement for written informed consent. Initially, 534 patients who underwent transrectal MRI-targeted biopsy using a dedicated MRI/ultrasound fusion system following multiparametric prostate MRI at our institution between June 2012 and March 2014 were identified. In patients who underwent more than one prostate MRI followed by fusion biopsy during this interval, only the first MRI and subsequent biopsy were considered for potential inclusion. The distribution of overall Likert scores on the prebiopsy MRI in this cohort was as follows: 2, 28.0% (150); 3, 31.6% (169); 4, 25.5% (136); 5, 14.8%.(79) As a score of 1 indicates an absence of suspicious findings, patients with this score would not be captured within a listing of patients undergoing fusion MRI-targeted biopsy. The 169 patients who received an overall Likert score of 3 were considered eligible for this study. Additional patients were then excluded for the following reasons: MRI performed at 1.5 Tesla (T) (n = 7); prior treatment for prostate cancer (n = 2); severe artifact from hip hardware (n = 2). These exclusions left a final included cohort of 158 patients (mean age, 64 ± 8 years). 44.3% (70/158) had no prior prostate biopsy, 28.5% (45/158) had a prior negative prostate biopsy, and 27.2% (43/158) had a prior positive prostate biopsy. The mean serum prostate-specific antigen was 6.0 ± 3.6 ng/mL. A total of 125 of these patients were included in a previous study comparing the diagnostic performance of cognitive and fusion biopsy approaches.[24] Table 1 summarizes characteristics of included patients.

TABLE 1. Characteristics of Study Cohort

Characteristic	Mean ± SD / frequency
Age	64±8 years
PSA	6.0±3.6 ng/mL
Biopsy history (by patient)	
Biopsy-naïve	44.3% (70/158)
Prior negative biopsy	28.5% (45/158)
Prior positive biopsy	27.2% (43/158)
Biopsy history (by lesion)	
Biopsy-naïve	42.6% (80/188)
Prior negative biopsy	28.2% (53/188)
Prior positive biopsy	29.3% (55/188)

MRI and MRI-Ultrasound Fusion Biopsy

All patients underwent multiparametric prostate MRI at 3T using an anterior pelvic phased-array coil in combination with a posterior spine coil array and a whole-body clinical system (Siemens MAGNETOM Trio, Skyra, or Verio). Sequences included axial turbo-spin echo (TSE) T2-weighted imaging (T2WI) (TR/TE 4000–4960/105 ms, slice thickness 3 mm, field-of-view (FOV) 180 × 180 mm, matrix 256 × 256, parallel imaging factor 2, 3 averages) and axial fat-suppressed single-shot echo-planar imaging diffusion-weighted imaging (DWI) (repetition time/echo time [TR/TE] 4100/86 ms, slice thickness 3 mm, FOV 200 × 200 mm, matrix 100 × 220, parallel imaging factor 2, 10 signal averages). DWI was performed using b-values of 50 and 1000 s/mm^2, with reconstruction of ADC maps as well as generation of computed DWI at a b-value of 1500 s/mm^2,[25] both performed using a standard monoexponential fit. In addition, dynamic contrast-enhanced (DCE) MRI was performed using an axial three-dimensional (3D) gradient-echo T1-weighted imaging (T1WI) sequence and intravenous administration of 0.1 mmol/kg of gadopentatate dimeglumine (Magnevist, Bayer

Healthcare Pharmaceuticals; n = 153) or, gadobutrol (Gadavist, Bayer Healthcare Pharmaceuticals; n = 5), which became the routine contrast agent used in our department in January 2014. In 107 patients, a conventional DCE sequence was performed (TR/TE 2.84–0.94 ms, flip angle 16°, slice thickness 3 mm, FOV 240 × 240 mm, matrix 128 × 128, no parallel imaging; 55 time-points at a temporal resolution of 5.5 s). In the remaining patients, a previously reported sequence that uses a golden-angle radial sampling scheme and parallel imaging reconstruction with compressed sensing was used (3192 radial spokes; TR/TE 4.10/1.89 ms, flip angle 16°, slice thickness 3 mm, FOV 240 × 240 mm, matrix 224 × 224; reconstructed with 21 radial spokes per time-point, providing 145 time-points at a temporal resolution of 2.3 s).[26]

A single fellowship-trained abdominal radiologist with 6 years of experience in prostate MRI (A.R.) reviewed all examinations and, prior to fusion biopsy, assigned a suspicion score to each detected lesion on a 2–5 Likert scale[12]; a score of 1 was reserved for a negative examination, without findings suspicious for tumor. All image sets (T2WI, high b-value DWI, ADC maps, DCE) were taken into account in assigning the Likert scores, although no specific ADC threshold was applied. While criteria for assessing each individual sequence that were previously proposed by an expert panel were considered,[11] the final lesion score was ultimately at the discretion of the radiologist and not determined by any specific rules. The highest score assigned to any individual lesion was considered to represent the overall score for the examination. Thus, a lesion scored as 3 in the setting of a concomitant score 4 lesion was not included in this analysis. This approach was selected given that a biopsy would be warranted in such a patient due to the presence of the score 4 lesion, regardless of the approach to the score 3 lesion.

Image Assessment

At the time of this retrospective study, the radiologist who had assigned the suspicion scores prior to fusion biopsy assessed the ADC map for each patient using in-house software (A.R.), blinded to subsequent biopsy results. The mean delay between the date of the MRI and date of this additional assessment was 465 ± 195 days (median, 414 days). Using this software, a 3D volume-of-interest (VOI) was placed encompassing the entire lesion volume on all slices depicting the lesion. No specific attempt was made to avoid the lesion edge. A histogram of ADC values within the lesion was generated based on the VOI, allowing for computation of whole-lesion ADC metrics: mean, kurtosis, skewness, and entropy of ADC. The mean ADC of adjacent portions of the histogram determined by the 10th, 25th, and 50th percentiles were also computed, hereafter referred to as $ADC_{0–10}$, $ADC_{10–25}$, and $ADC_{25–50}$. These metrics provide a more robust assessment of the presence of low ADC values within the lesion than does the absolute minimum ADC value within any single voxel, which is more strongly influenced by random variation.[22] The volume of each VOI was also computed.

For purposes of assessing inter-reader agreement, a second radiologist with 2 years of experience in interpretation of prostate MRI (J.R.) placed VOIs using the same software and also recorded the whole-lesion histogram metrics. This reader was aware of the general location of the score 3 lesions at the time of prospective interpretation, but blinded to the VOIs placed by the first reader.

Reference Standard

Transrectal biopsy in all patients was guided by the ei-Nav|Artemis system (Eigen, Grass Valley, CA), with the ProFuse software (Eigen) used for prostate segmentation and marking of targets. Using this system, MRI-ultrasound fusion was used to obtain targeted cores from all lesions reported on MRI. Standard systematic 12-core biopsy was also performed during the same session. Given multiple investigations estimating a registration error of 3.5 mm of current fusion systems,[27–29] each score 3 lesion on MRI was considered to correspond with tumor if tumor was detected either in the targeted cores or in the nearest adjacent systematic core (n = 10). Detected tumors were classified as GS = 6 or GS > 6 for purposes of analysis.

Statistical Assessment

Logistic regression for correlated data was used to assess the association of clinical parameters (patient age, prostate specific antigen [PSA], previous biopsy history), lesion volume on the ADC map, and the whole-lesion ADC metrics with both the presence of tumor and of GS > 6 tumor. Specifically, generalized estimating equations based on logistic regression was used to model each of these two outcomes as a function of each of the candidate predictors while accounting for within-subject correlations due to the inclusion of results from two or three score 3 lesions within some patients. The analysis assumed outcomes to be correlated only when derived for lesions within the same patient. For the purposes of analysis, the patient's biopsy history was assigned to one of three categories (biopsy-naïve, prior negative biopsy, prior positive biopsy) and then applied in an ordinal manner within the regression model based on the frequency of tumor observed within the three cohorts. For all metrics that were significant predictors of GS > 6 tumor, thresholds were identified to detect GS > 6 tumor in the various patient categories at a target sensitivity of 90%, based on the method of Zhou et al.[30] The single-measures intraclass correlation coefficient (ICC) for absolute agreement was used to assess inter-reader agreement of the histogram metrics and classified as follows: <0.2, no correlation; between 0.2 and 0.4 weak correlation; between 0.4 and 0.6, moderate correlation; between 0.6 and 0.8, good correlation; over 0.8, excellent correlation.[31] All statistical tests were conducted at the two-sided 5% significance level using SAS Version 9.3 (SAS Institute, Cary, NC) and MedCalc for Windows Version 12.7 (MedCalc Software, Ostend, Belgium).

Results

Patients

A total of 188 Likert score 3 lesions were present in the 158 patients. Based on subsequent biopsy results, 22.3% (43/188) were classified as positive for tumor, and 8.5% (16/188) as positive for GS > 6 tumor. Among the GS > 6 tumors, two were Gleason score 4+3 and one was Gleason score 4+4; the remainder were Gleason score 3+4. By lesion, 78.2% (147/188) were in the peripheral zone, of which 24.5% (36/147) were positive for tumor and 8.8%

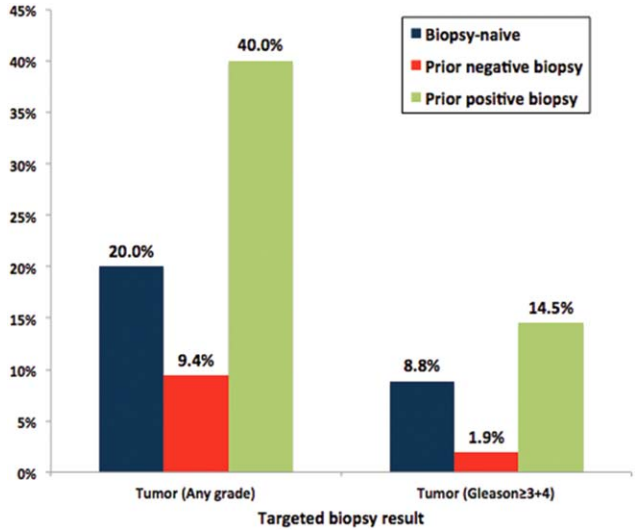

FIGURE 1: Graph comparing frequency of any tumor and of Gleason score >6 tumor identified at MRI/ultrasound fusion targeted biopsy of Likert score 3 lesions, stratified by patient biopsy history.

(13/147) were positive for GS > 6 tumor. The remaining 21.8% (41/188) of lesions were in either the transition or central zone, of which 17.1% (7/41) were positive for tumor and 7.3% (3/41) were positive for GS > 6 tumor. Among the 43 patients with a prior positive biopsy, the Gleason score on prior biopsy was 3+3 in 39 patients and 3+4 in 4 patients; 14.0% (6/43) of these patients were upgraded based on the results of the MRI/ultrasound fusion biopsy.

Figure 1 shows that both the frequency of any tumor ($P < 0.001$) and of GS > 6 tumor ($P = 0.017$) varied based on patient biopsy history, being lowest in patients with a prior negative biopsy, intermediate in biopsy-naïve patients, and highest in patients with a prior positive biopsy. Additional factors that were significantly associated with presence of tumor were mean ADC, $ADC_{10–25}$, and $ADC_{25–50}$ ($P = 0.041–0.047$) (Table 2). These three factors, along with PSA, were also significantly associated with presence of GS > 6 tumor ($P = 0.008–0.034$). Patient age, lesion volume, $ADC_{0–10}$, skewness ADC, kurtosis ADC, and entropy ADC were not significant predictors of either of these pathologic outcomes ($P \geq 0.071$). A representative patient is demonstrated in Figure 2.

Table 3 shows the performance of various metrics in detection of GS > 6 tumor at thresholds identified to achieve 90% sensitivity. Only biopsy-naïve patients as well as patients with a prior positive biopsy are included in this analysis given the presence of GS > 6 tumor in only one patient with a prior negative biopsy. In the patients without a prior negative biopsy, $ADC_{10–25}$ below 1.05×10^{-3} mm^2/s identified GS > 6 tumor with a specificity of 50.0% at 90% sensitivity. This specificity was significantly higher ($P < 0.001–0.006$) than that of PSA (specificity of 17.5%) and of $ADC_{25–50}$ (specificity of 41.7%), although not significantly higher ($P = 0.549$) than that of mean ADC (specificity of 47.5%).

Inter-reader agreement was excellent for mean ADC, $ADC_{10–25}$, and $ADC_{25–50}$; good for $ADC_{0–10}$ and entropy; and weak for skewness and kurtosis (Table 4).

TABLE 2. Association of Study Variables (Listed as Mean ± SD) With Pathologic Outcomes at Fusion Biopsy of Score 3 Lesions[a]

Measure	Any tumor			GS > 6 tumor		
	Yes	No	*P*	Yes	No	*P*
Biopsy history	[b]		**<0.001**	[b]		**0.017**
Age (years)	64.5±8.5	63.4±7.8	0.480	61.4±10.4	63.8±7.7	0.279
PSA (ng/mL)	5.6±2.6	6.1±3.9	0.402	4.8±1.9	6.1±3.7	**0.034**
Lesion volume (cm^3)	0.27±0.28	0.29±0.42	0.682	0.21±0.13	0.30±0.41	0.172
Histogram metrics						
Entropy	3.95±0.89	3.94±0.84	0.895	3.93±0.64	3.94±0.87	0.863
Mean ADC[c]	1.08±0.19	1.15±0.20	**0.041**	1.02±0.15	1.14±0.20	**0.008**
$ADC_{0–10}$[c]	0.85±0.23	0.92±0.25	0.103	0.83±0.15	0.91±0.26	0.071
$ADC_{10–25}$[c]	0.95±0.19	1.01±0.21	**0.047**	0.91±0.15	1.01±0.21	**0.019**
$ADC_{25–50}$[c]	1.03±0.19	1.09±0.21	**0.047**	0.97±0.15	1.09±0.21	**0.008**
Kurtosis	−0.04±0.98	−0.03±1.19	0.985	0.06±0.74	−0.04±1.18	0.756
Skewness	0.25±0.48	0.23±0.55	0.803	0.22±0.45	0.24±0.54	0.803

[a]Bold highlighting indicates statistically significant at the 0.05 significance level.
[b]See Figure 1.
[c]$\times 10^{-3}$ mm^2/s

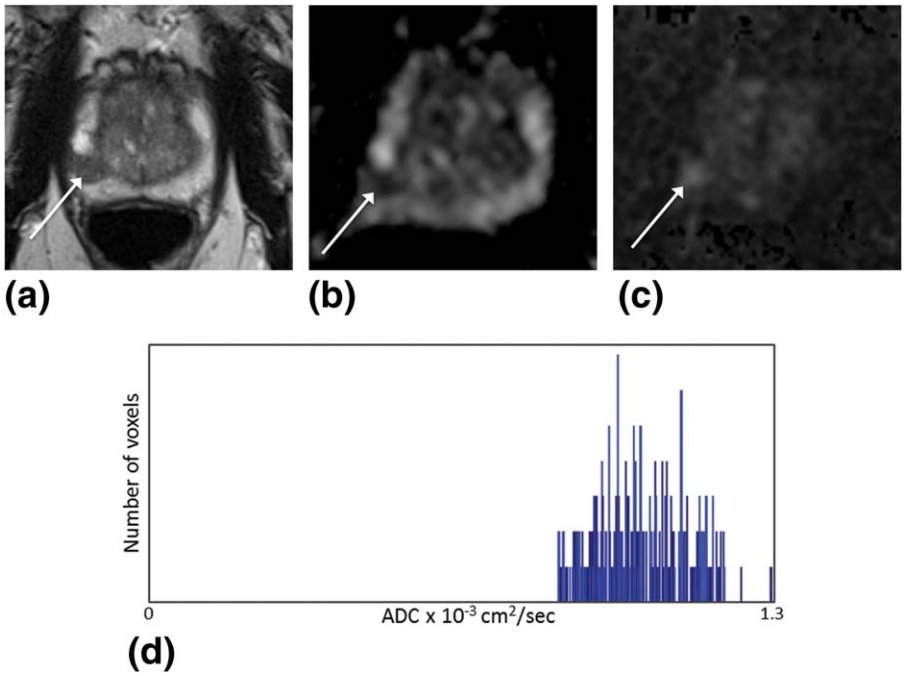

FIGURE 2: A 70-year-old man with prior negative prostate biopsy and persistently elevated PSA of 8.67 ng/mL. A: Axial turbo spin-echo T2-weighted image (T2WI) demonstrates a geographic region of mildly decreased T2 signal (arrow). B: Axial apparent diffusion coefficient (ADC) map shows corresponding mildly decreased ADC (arrow). C: Computed diffusion-weighted image with a b-value of 1500 s/mm^2 shows mildly increased signal intensity (arrow). Lesion was categorized as Likert score 3 given the mild reduction in T2 signal intensity and ADC, as well as its geographic, non–mass-like shape. D: Histogram of ADC values obtained from whole-lesion volume-of-interest on all slices encompassing lesion on ADC map; based on the histogram, lesion exhibited mean ADC of 1.30 × 10^{-3} mm^2/s. Lesion was benign on MRI/ultrasound fusion targeted biopsy. Follow-up PSA a year later was slightly decreased at 7.4 ng/mL.

Discussion

While targeted biopsy is advised in the setting of a suspicion score of 4 or 5 on prostate MRI, and deferral of biopsy may be considered for a score of 1 or 2, the optimal management of score 3 lesions has remained uncertain. We observed significant associations between both patient biopsy history and ADC-based metrics with pathologic outcomes of Likert score 3 prostate lesions undergoing MRI/ultrasound fusion targeted biopsy. In general, the overall frequency of tumor within the study cohort was low, which is reassuring given past suggestions to routinely target lesions receiving a score of 4 or 5. Nonetheless, the frequency of GS > 6 tumors within the cohort was 8.5%. Thus, routinely deferring biopsy of score 3 lesions risks missing a considerable number of aggressive tumors when applied across large patient populations.

The patient's biopsy history was strongly associated with targeted biopsy results, both in terms of overall tumor and of GS > 6 tumor. In particular, the positivity rate for these outcomes was highest in patients with a prior positive

TABLE 3. Specificity of Metrics in Identification of Gleason Score > 6 Tumor at Fusion Biopsy of Likert Score 3 Lesions in Patients Without a Prior Negative Biopsy[a]

Parameter	Threshold	Specificity
PSA	≤ 7.15	17.5% (21/120)
Mean ADC[b]	≤ 1.18	47.5% (57/120)
$ADC_{10–25}$[b]	≤ 1.05	50.0% (60/120)
$ADC_{25–50}$[b]	≤ 1.14	41.7% (50/120)

[a]Thresholds are selected to achieve 90% sensitivity for Gleason score > 6 tumor.
[b]×10^{-3} mm^2/s.

TABLE 4. Intraclass Correlation Coefficient for Whole-Lesion ADC Metrics

Metric	ICC
Mean ADC	0.81
$ADC_{0–10}$	0.63
$ADC_{10–25}$	0.85
$ADC_{25–50}$	0.85
Skewness	0.28
Kurtosis	0.24
Entropy	0.74

biopsy, lowest in patients with a prior negative biopsy, and intermediate in biopsy-naïve patients. These differences are not unexpected given the known ability of benign conditions, such as prostatitis, high-grade prostatic intraepithelial neoplasia, and benign prostate hyperplasia, to mimic tumor on MRI,[32] and the greater likelihood of a detected lesion to in fact represent tumor when this diagnosis has previously been established. Therefore, while a score of 3 should be assigned strictly based on the imaging findings in a given case, and not raised or lowered on account of the biopsy history, it is important for the biopsy history to be taken into consideration when deciding whether to perform a targeted biopsy of such a lesion.

One recent study of transperineal in-bore MRI-guided biopsy also observed an impact of biopsy history on targeted biopsy outcomes.[33] In that study, 48.1% of patients without previously diagnosed prostate cancer, and 61.5% of patients with known prostate cancer, were positive at MRI-targeted biopsy.[33] That study did not differentiate between biopsy-naïve patients and patients with a prior negative biopsy. In addition, biopsy results were not stratified by MRI suspicion score, such that the findings were influenced by the inclusion of high suspicion MRI lesions that typically warrant targeted biopsy, regardless of patient history.

Our observations may help in selection of patients with a score 3 lesion for targeted biopsy. Given the exceedingly low frequency of GS > 6 tumor in Likert score 3 lesions in patients with a prior negative biopsy (approximately 1 out of 50 such patients), we suggest that deferral of targeted biopsy may be considered for such lesions. Although biopsy of Likert score 3 lesions should be considered in remaining patients, targeting all such lesions in patients without a prior negative biopsy would result in a high fraction of benign biopsy results. Rather, the ADC metrics can be used to reduce the number of negative biopsies in this context. When considering threshold values to apply in practice, high sensitivity for significant tumor is desired. Among patients without a prior negative biopsy, selecting score 3 lesions for biopsy based on $ADC_{10–25} \leq 1.05 \times 10^{-3}$ mm^2/s would have avoided biopsy in approximately half of lesions that were either benign or low-grade tumor, while maintaining 90% sensitivity for GS > 6 tumor. Therefore, while the initial stratification of MRI findings using the 1–5 scale already facilitates reducing the number of biopsies by avoiding biopsy in the setting of a score of 1 or 2, the ADC metrics have a role for substantially reducing the number of biopsies further. Although the specificity at the derived threshold is low, this specificity was significantly higher than that of PSA, which currently serves as the primary noninvasive marker to guide biopsy decisions.

The 1–5 MRI suspicion score that currently guides decisions regarding targeted biopsy is based on visual image assessment and does not consider quantitative metrics. Our observations support the complimentary role of ADC metrics to further refine clinical algorithms and more precisely select individual patients within a given cohort for biopsy. The known inverse correlation between ADC values and Gleason score,[18] as well as the inherent invisibility on ADC maps of a fraction of GS = 6 tumors,[34] presumably contributed to the particularly strong association of the ADC metrics with GS > 6 tumor, in comparison with these metrics' association with overall tumor. Indeed, any associations that we observed between the whole-lesion ADC metrics and presence of any tumor are of questionable relevance once considering multiple comparisons testing and the obtained levels of statistical significance for these assessments. Rather, our findings suggest that the utility of the whole-lesion ADC metrics is largely for the detection of GS > 6 tumors. It is notable that the frequency of GS > 6 tumor was overall low within the cohort, regardless of biopsy history. Such lesions are expected to typically receive a suspicion score of 4 or 5, and would therefore be anticipated to be uncommon within this current investigation of score 3 lesions.

Our analysis included histogram-based ADC metrics derived from whole-lesion assessment. Such metrics are intended to better assess lesion texture and have shown added value compared with traditional metrics such as mean or median ADC in past studies.[21–23] In this study, these various measures that reflect the shape and variation of the histogram, such as skewness, kurtosis, and entropy, failed to exhibit significant associations with biopsy results. We speculate that this may relate to the generally small size of the lesions in the current study. These textural metrics reflect the distribution of ADC values within the histogram. Differences in such distributions may be less apparent for histograms derived from a relatively smaller number of voxels. We do note that $ADC_{10–25}$ and $ADC_{25–50}$ were significantly associated with both pathologic outcomes. The value of these metrics assessing lower ADC regions within the histogram may relate to the histologic composition of prostate cancer, exhibiting intermixed benign and malignant glandular elements.[34] Metrics assessing lower ADC may be less influenced by such intermixed benign regions and better reflect focal malignant components within the overall lesion volume.[21] Furthermore, while Likert 3 lesions may be difficult to precisely delineate on the ADC maps in comparison with higher suspicion lesions, excellent interobserver reproducibility was observed for the metrics that yielded highest predictive value in this study. Although some variability between readers in the boundaries of the region of interest (ROI) is anticipated, our findings suggest that such variability did not substantially alter the percentile-based ADC metrics between readers. In comparison, the texture based-metrics (skewness and kurtosis) exhibited poor inter-reader agreement, presumably related to the impact of slight variation in ROI boundaries upon overall lesion texture,

particularly in the context of small lesions as were common in our cohort.

This study has several limitations. First, suspicion scores were assigned on the basis of a Likert scale, not entailing fixed criteria for each numeric value. However, a recent study in a large patient population undergoing radical prostatectomy observed the Likert score to be significantly more accurate in categorization of prostate lesions than other scoring schemes, including PI-RADS.[9] An additional study observed similar performance of tumor localization using PI-RADS and Likert scales, although in the transition zone, performance was better using the Likert score.[12] Furthermore, at the time of this writing, the PI-RADS scheme is continuing to undergo optimization and development.[10] Therefore, we believe that the Likert scale provides an appropriate classification scheme for investigation in this study. Second, all Likert scales were assigned prospectively by a single radiologist. However, prior work suggested excellent inter-reader reproducibility in the peripheral zone and moderate interobserver reproducibility in the TZ for Likert scores assigned by experienced observers from a single institution.[35] Third, targeted biopsy served as reference standard, whereas pathologic findings from radical prostatectomy would be more accurate. However, this approach is reasonable given our focus on identification of important tumors on targeted biopsy and patient selection for this procedure. In addition, radical prostatectomy would not be performed on patients with negative biopsy, as well as in a substantial fraction of patients with low-grade tumor on biopsy, such that it would not have been possible to use prostatectomy findings as the reference standard in this study. Fourth, a fraction of lesions were classified as positive based on tumor detection in the nearest adjacent sextant. However, phantom studies suggest a registration error of 3–4 mm for current fusion technology,[27] such that classifying all such instances as benign would likely comprise a false-negative classification for some lesions. Novel fusion approaches are under development that may improve the accuracy of MRI/ultrasound registration in the future and facilitate more precise comparisons of fusion biopsy findings with pathology outcomes.[29] Fifth, we did not explore the role of changes over time in an MRI lesion in men who had undergone serial MRI evaluation before fusion biopsy. A past study suggests that lesion progression on MRI also serves as a strong predictor of significant cancer.[36] Sixth, while the suspicion scores were assigned prospectively at the time of original interpretation, the whole-lesion histogram metrics were derived retrospectively following the time of fusion biopsy. Thus, the performance of this quantitative approach in clinical practice is unknown. Seventh, this study was performed at a major referral center for prostate cancer management, possibly resulting in an overall selection bias. It therefore is important to confirm the observations in other practice settings. Eighth, as previously noted, we did not account for multiple comparisons in our statistical analyses. Finally, an endorectal coil was not used. However, current consensus guidelines indicate that an endorectal coil is not needed for tumor localization, which was the primary focus of this investigation as well as of targeted fusion biopsy. In addition, a recent survey indicated that the most widely used protocol for prostate MRI at academic medical centers in the United States comprises imaging at 3T without an endorectal coil,[37] thus matching the approach used in our analysis. One disadvantage of not using the endorectal coil in this particular study was potential anatomic distortion of the ADC maps due to artifact from rectal gas. Nonetheless, we demonstrated utility of the ADC metrics despite not excluding any patients on this basis.

In conclusion, for detection of $GS > 6$ tumor, we suggest that deferral of targeted biopsy of score 3 lesions can be considered, without application of whole-lesion ADC metrics, in patients with a prior negative biopsy given the exceedingly low frequency of this outcome in this cohort (approximately 1 out of 50 patients). In other patient groups (in whom the likelihood of $GS > 6$ tumor in score 3 lesions was higher), targeted biopsy of score 3 lesions with a whole-lesion $ADC_{10-25} \leq 1.05 \times 10^{-3}$ mm^2/s provides 90% sensitivity and 50% specificity for detection of $GS > 6$ tumor, thereby helping to reduce the frequency of negative biopsies compared with use of PSA as the selection criteria for biopsy. These observations can be used to help guide selection of patients with a score 3 lesion on multiparametric prostate MRI for targeted biopsy.

Acknowledgments

Contract grant sponsor: The Joseph and Diane Steinberg Charitable Trust.

Author Samir S. Taneja is a consultant for Hitachi-Aloka and for Healthtronics; receives payments for lectures as well as travel/accommodation expenses from Hitachi-Aloka; and receives royalties from Elsevier. None of the remaining authors have any disclosures to report.

References

1. Hoeks CM, Barentsz JO, Hambrock T, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. Radiology 2011;261:46–66.
2. Hoeks CM, Somford DM, van Oort IM, et al. Value of 3-T multiparametric magnetic resonance imaging and magnetic resonance-guided biopsy for early risk restratification in active surveillance of low-risk prostate cancer: a prospective multicenter cohort study. Invest Radiol 2014;49:165–172.
3. Siddiqui MM, Rais-Bahrami S, Truong H, et al. Magnetic resonance imaging/ultrasound-fusion biopsy significantly upgrades prostate cancer versus systematic 12-core transrectal ultrasound biopsy. Eur Urol 2013;64:713–719.

4. Vourganti S, Rastinehad A, Yerram NK, et al. Multiparametric magnetic resonance imaging and ultrasound fusion biopsy detect prostate cancer in patients with prior negative transrectal ultrasound biopsies. J Urol 2012;188:2152–2157.
5. Kaufmann S, Kruck S, Kramer U, et al. Direct comparison of targeted MRI-guided biopsy with systematic transrectal ultrasound-guided biopsy in patients with previous negative prostate biopsies. Urol Int 2015;94:319–325.
6. Stamatakis L, Siddiqui MM, Nix JW, et al. Accuracy of multiparametric magnetic resonance imaging in confirming eligibility for active surveillance for men with prostate cancer. Cancer 2013;119:3359–3366.
7. Loch R, Fowler K, Schmidt R, Ippolito J, Siegel C, Narra V. Prostate magnetic resonance imaging: challenges of implementation. Curr Probl Diagn Radiol 2015;44:26–37.
8. Dickinson L, Ahmed HU, Allen C, et al. Scoring systems used for the interpretation and reporting of multiparametric MRI for prostate cancer detection, localization, and characterization: could standardization lead to improved utilization of imaging within the diagnostic pathway? J Magn Reson Imaging 2013;37:48–58.
9. Vache T, Bratan F, Mege-Lechevallier F, Roche S, Rabilloud M, Rouviere O. Characterization of prostate lesions as benign or malignant at multiparametric MR imaging: comparison of three scoring systems in patients treated with radical prostatectomy. Radiology 2014; 272:446–455.
10. Bomers JG, Barentsz JO. Standardization of multiparametric prostate MR imaging using PI-RADS. BioMed Res Int 2014;2014:431680.
11. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. Eur Radiol 2012;22:746–757.
12. Rosenkrantz AB, Kim S, Lim RP, et al. Prostate cancer localization using multiparametric MR imaging: comparison of Prostate Imaging Reporting and Data System (PI-RADS) and Likert scales. Radiology 2013;269:482–492.
13. Junker D, Quentin M, Nagele U, et al. Evaluation of the PI-RADS scoring system for mpMRI of the prostate: a whole-mount step-section analysis. World J Urol 2014. [Epub ahead of print].
14. Vargas HA, Akin O, Afaq A, et al. Magnetic resonance imaging for predicting prostate biopsy findings in patients considered for active surveillance of clinically low risk prostate cancer. J Urol 2012;188: 1732–1738.
15. Yerram NK, Volkin D, Turkbey B, et al. Low suspicion lesions on multiparametric magnetic resonance imaging predict for the absence of high-risk prostate cancer. BJU Int 2012;110(Pt B):E783–E788.
16. Roethke MC, Kuru TH, Schultze S, et al. Evaluation of the ESUR PI-RADS scoring system for multiparametric MRI of the prostate with targeted MR/TRUS fusion-guided biopsy at 3.0 Tesla. Eur Radiol 2014; 24:344–352.
17. Grey AD, Chana MS, Popert R, Wolfe K, Liyanage SH, Acher PL. Diagnostic accuracy of magnetic resonance imaging (MRI) prostate imaging reporting and data system (PI-RADS) scoring in a transperineal prostate biopsy setting. BJU Int 2015;115:728–735.
18. Hambrock T, Somford DM, Huisman HJ, et al. Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. Radiology 2011;259:453–461.
19. Park JJ, Kim CK, Park SY, Park BK, Lee HM, Cho SW. Prostate cancer: role of pretreatment multiparametric 3-T MRI in predicting biochemical recurrence after radical prostatectomy. AJR Am J Roentgenol 2014;202:W459–W465.
20. Giles SL, Morgan VA, Riches SF, Thomas K, Parker C, deSouza NM. Apparent diffusion coefficient as a predictive biomarker of prostate cancer progression: value of fast and slow diffusion components. AJR Am J Roentgenol 2011;196:586–591.
21. Donati OF, Mazaheri Y, Afaq A, et al. Prostate cancer aggressiveness: assessment with whole-lesion histogram analysis of the apparent diffusion coefficient. Radiology 2014;271:143–152.
22. Peng Y, Jiang Y, Yang C, et al. Quantitative analysis of multiparametric prostate MR images: differentiation between prostate cancer and normal tissue and correlation with Gleason score–a computer-aided diagnosis development study. Radiology 2013;267:787–796.
23. Rosenkrantz AB, Triolo MJ, Melamed J, Rusinek H, Taneja SS, Deng FM. Whole-lesion apparent diffusion coefficient metrics as a marker of percentage Gleason 4 component within Gleason 7 prostate cancer at radical prostatectomy. J Magn Reson Imaging 2015;41:708–714.
24. Wysock JS, Rosenkrantz AB, Huang WC, et al. A prospective, blinded comparison of magnetic resonance (MR) imaging-ultrasound fusion and visual estimation in the performance of MR-targeted prostate biopsy: the PROFUS Trial. Eur Urol 2014;66:343–351.
25. Rosenkrantz AB, Chandarana H, Hindman N, et al. Computed diffusion-weighted imaging of the prostate at 3 T: impact on image quality and tumour detection. Eur Radiol 2013;23:3170–3177.
26. Rosenkrantz AB, Geppert C, Grimm R, et al. Dynamic contrast-enhanced MRI of the prostate with high spatiotemporal resolution using compressed sensing, parallel imaging, and continuous golden-angle radial sampling: preliminary experience. J Magn Reson Imaging 2015;41:1365–1373.
27. Martin PR, Cool DW, Romagnoli C, Fenster A, Ward AD. Magnetic resonance imaging-targeted, 3D transrectal ultrasound-guided fusion biopsy for prostate cancer: quantifying the impact of needle delivery error on diagnosis. Med Phys 2014;41:073504.
28. Sparks R, Bloch BN, Feleppa E, et al. Multiattribute probabilistic prostate elastic registration (MAPPER): application to fusion of ultrasound and magnetic resonance imaging. Med Phys 2015;42:1153–1163.
29. Fedorov A, Khallaghi S, Sanchez CA, et al. Open-source image registration for MRI-TRUS fusion-guided prostate interventions. Int J Comput Assist Radiol Surg 2015;10:925–934.
30. Zhou XH, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. 2nd ed. New York: Wiley Interscience; 2011.
31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
32. Rosenkrantz AB, Taneja SS. Radiologist, be aware: ten pitfalls that confound the interpretation of multiparametric prostate MRI. AJR Am J Roentgenol 2014;202:109–120.
33. Penzkofer T, Tuncali K, Fedorov A, et al. Transperineal in-bore 3-T MR imaging-guided prostate biopsy: a prospective clinical observational study. Radiology 2015;274:170–180.
34. Langer DL, van der Kwast TH, Evans AJ, et al. Intermixed normal tissue within prostate cancer: effect on MR imaging measurements of apparent diffusion coefficient and T2–sparse versus dense cancers. Radiology 2008;249:900–908.
35. Rosenkrantz AB, Lim RP, Haghighi M, Somberg MB, Babb JS, Taneja SS. Comparison of interreader reproducibility of the prostate imaging reporting and data system and likert scales for evaluation of multiparametric prostate MRI. AJR Am J Roentgenol 2013;201:W612–W618.
36. Rosenkrantz AB, Rice SL, Wehrli NE, Deng FM, Taneja SS. Association between changes in suspicious prostate lesions on serial MRI examinations and follow-up biopsy results. Clin Imaging 2015;39:264–269.
37. Leake JL, Hardman R, Ojili V, et al. Prostate MRI: access to and current practice of prostate MRI in the United States. J Am Coll Radiol 2014;11:156–160.