Published in final edited form as:

AJR Am J Roentgenol. 2019 April ; 212(4): 823–829. doi:10.2214/AJR.18.20295.

The Influence of Background Signal Intensity Changes on Cancer Detection in Prostate MRI

Andreas M. Hötter¹, Evelyn Dappa², Yousef Mazaheri³, Behfar Ehdaie^{4,5}, Junting Zheng⁵, Marinela Capanu⁵, Hedvig Hricak⁶, and Oguz Akin⁶

¹Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland. ²Department of Diagnostic and Interventional Radiology, Universitätsmedizin Mainz, Mainz, Germany. ³Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY. ⁴Department of Surgery, Urology Service, Memorial Sloan-Kettering Cancer Center, New York, NY. ⁵Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY. ⁶Department of Radiology, Memorial Sloan-Kettering Cancer Center, New York, NY.

Abstract

OBJECTIVE.—The objective of this study was to develop a scoring system for background signal intensity changes or prostate homogeneity on prostate MRI and to assess these changes' influence on cancer detection.

MATERIALS AND METHODS.—This institutional review board–approved, HIPAA-compliant, retrospective study included 418 prostate MRI examinations in 385 men who subsequently underwent MRI-guided biopsy. The Likert score for suspicion of cancer assigned by the primary radiologist was extracted from the original report, and histopathologic work-up of the biopsy cores served as the reference standard. Two readers assessed the amount of changes on T2-weighted sequences and assigned a predefined prostate signal-intensity homogeneity score of 1–5 (1 = poor, extensive changes; 5 = excellent, no changes). The sensitivity and specificity of Likert scores for detection of prostate cancer and clinically significant cancer (Gleason score $\geq 3+4$) were estimated in and compared between subgroups of patients with different signal-intensity homogeneity scores (2, 3, and 4).

RESULTS.—Interreader agreement on signal-intensity homogeneity scores was substantial ($\kappa = 0.783$). Sensitivity for prostate cancer detection increased when scores were better (i.e., higher) (reader 1, from 0.41 to 0.71; reader 2, from 0.53 to 0.73; $p = 0.007$, both readers). In the detection of significant cancer (Gleason score $\geq 3+4$), sensitivity also increased with higher signal-intensity scores (reader 1, from 0.50 to 0.82; reader 2, from 0.63 to 0.86; $p = 0.028$), though specificity decreased significantly for one reader (from 0.67 to 0.38; $p = 0.009$).

CONCLUSION.—Background signal-intensity changes on T2-weighted images significantly limit prostate cancer detection. The proposed scoring system could improve the standardization of

prostate MRI reporting and provide guidance for applying prostate MRI results appropriately in clinical decision-making.

Keywords

MRI; prostate cancer

Multiparametric MRI (mpMRI) of the prostate has emerged as the imaging modality of choice for the detection and staging of prostate cancer. It has been found to improve cancer detection rates in patients with elevated prostate-specific antigen levels who undergo subsequent MRI-guided transrectal ultra-sound (TRUS) biopsy [1, 2]. The chance of identifying clinically significant cancer, in particular, increases if MRI is performed before biopsy to identify lesions that warrant targeting [2, 3]. However, identification of target lesions can be hindered by changes in the normal background signal intensity of the prostatic tissue, such as those associated with prostatitis, scarring, or postbiopsy hemorrhage. These changes appear as patchy areas of low signal intensity on standard T2-weighted images and can mimic tumors or obfuscate lesions that would otherwise be detected [4]. Although the appearance of postbiopsy changes can be mitigated by waiting a substantial period of time after biopsy to perform MRI, changes caused by inflammation or fibrous tissue are common and may be seen in both biopsy-naïve patients and those who undergo MRI a substantial period of time after biopsy [5].

In the interpretation of other imaging examinations used to detect cancer, comparable factors that hinder tumor detection are routinely assessed and noted in the radiologic report. For example, because dense breast tissue is known to obfuscate a significant number of tumors, grading of breast density is mandatory for all mammography reports and a legal requirement in some countries [6]. However, this paradigm does not exist in prostate MRI, and changes of the prostate background signal intensity are not categorized or classified in a standardized manner. In addition, to our knowledge the influence of these changes on diagnostic performance (i.e., cancer detection rates with MRI), has not been comprehensively assessed.

The purpose of this study was therefore to develop a standardized scoring system for prostate signal-intensity homogeneity on routine prostate MRI examinations and to investigate the influence of changes in homogeneity on cancer detection rates in patients undergoing MRI-guided biopsy.

Materials and Methods

Patients

This study was approved by the institutional review board at Memorial Sloan-Kettering Cancer Center, and the requirement for informed consent was waived. This retrospective study was compliant with HIPAA. We searched urology databases for the period from January 2014 through September 2016 to identify patients who underwent prostate mpMRI followed by MRI-guided TRUS biopsy. This search identified 397 patients, of whom eight were excluded because of bad image quality or severe motion artifacts, three because of prior radiotherapy, and one because of previous ablation (Fig. 1).

The final study cohort consisted of 385 men (mean age, 63 years old; range, 41–79 years old) for whom a total of 418 prebiopsy MRI examinations were included. Of these 385 patients, 16 underwent sequential MRI and biopsies in an active surveillance regimen (median time between biopsies, 643 days; range, 304–773 days). Clinical and histopathologic information was gathered from the patients' medical records. For each MRI examination, the prebiopsy MRI score for the presence of cancer was recorded from the original radiology report. This score was given by a board-certified radiologist with dedicated training in reading prostate MRI who read the scan for clinical purposes and was expressed on a 5-point Likert scale. A score of 1 indicated that the prostate cancer was highly unlikely; a score of 5 indicated that prostate cancer was highly likely. For statistical analysis, Likert scores of 4 and 5 were deemed to indicate positivity for cancer.

The biopsy operator had access to the radiology report as well as images that were previously interpreted and annotated for targeted biopsy by the reporting radiologist. Patients with a suspicious lesion (Likert score ≥ 3) on MRI underwent a targeted biopsy of each ROI, consisting of two cores obtained under visual registration and two cores obtained under software registration using a computer-assisted elastic MRI-ultrasound image fusion system with real-time 3D tracking technology (UroStation, Koelis). Patients then underwent an additional 14-core systematic TRUS-guided prostate biopsy consisting of samples obtained from the medial and lateral aspects of the base, middle, and apical portions of the prostate bilaterally, along with two samples from the transition zone excluding areas that had been sampled during the MRI-targeted biopsy. Patients with no suspicious lesion identified on MRI underwent systematic biopsy only.

MRI Technique

MRI was performed at field strengths of 3 T ($n = 407$) and 1.5 T ($n = 11$). An endorectal coil was used in 82 patients. All examinations adhered to a dedicated prostate MRI protocol, including axial, coronal, and sagittal T2-weighted turbo spin-echo sequences (TR/TE, 2300–8682/88–229; acquisition matrix, 320×224 –256; FOV, 180–220 mm $\times$ 180–220 mm; slice thickness, 3–4 mm), as well as a transverse T1-weighted sequence, a diffusion-weighted sequence (b values of 0, 400, and 1000 s/mm² or 0, 700, and 1000 s/mm²), and a T1-weighted dynamic contrast-enhanced sequence covering the pelvis in 206 patients.

Development of Scoring System

The scoring system was developed on the basis of expert consensus during our institution's research and disease management team meetings. It consists of a scale from 1 to 5, indicating the amount of T2 hypointense changes and grade of homogeneity in the peripheral zone. A low score signifies an amount of changes in the peripheral zone great enough to markedly hinder the assessment of focal prostatic lesions, and a score of 5 indicates a homogeneous, T2-hyperintense peripheral zone that does not hinder lesion assessment (Fig. 2).

Two readers, blinded to all clinical and histopathologic information, independently reviewed all scans retrospectively. Each assigned a signal homogeneity score based on the scoring system illustrated in Figures 2 and 3. Areas clearly suspicious for cancer (i.e., nodular

hypointensity on T2-weighted images with diffusion restriction and early contrast material uptake with an initial Likert score of 4 or 5) were not considered when deriving the score.

Statistical Analysis

To examine differences in clinical factors and biopsy outcomes with different prostate signal-intensity homogeneity scores (2 vs 3 vs 4), univariate multinomial regressions were used with a cumulative logit link function. Because some patients (16/385) underwent MRI more than once, the generalized estimating equations (GEE) method was used with an independent correlation matrix and an empirical covariance matrix. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for the detection of prostate cancer or clinically significant cancer (Gleason score 3+4) were estimated in the three subgroups of patients with different signal-intensity homogeneity scores. To compare sensitivities, specificities, PPVs, and NPVs between these subgroups, generalized logistic regressions and the GEE method were used. These analyses were performed for each reader separately.

The weighted kappa with squared weights was used to assess agreement on signal-intensity homogeneity scores between the two readers. The CI of the weighted kappa was estimated using bootstrap with 2000 repetitions. The kappa value was interpreted as follows: 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement.

A test with a p value less than 0.05 was considered statistically significant. All statistical analyses were performed using SAS (version 9.4, SAS Institute) and R (version 3.3, R Foundation for Statistical Computing) software packages.

Results

Patient and Tumor Characteristics

Detailed patient characteristics and biopsy results are given in Table 1. The proportions of cancer and significant cancer (i.e., Gleason score 3 + 4) were not different among signal-intensity homogeneity score groups. The total number of cores, the percentage of cancer, and the length of cancer in the most aggressive core did not differ by subgroup either. However, we found a statistically significant difference in Likert scores assigned by primary radiologists between the subgroups ($p < 0.001$ for both readers).

Interreader Agreement

The two readers showed substantial agreement on signal-intensity homogeneity scores (Table 2). The estimated weighted kappa reached 0.783 (95% CI, 0.733–0.818), and concordant readings were seen in 280/418 (67%) of all examinations.

Influence of Background Signal Intensity Changes on Cancer Detection Rates

The overall sensitivity and specificity of prebiopsy MRI Likert scores were 0.74 (113/153) and 0.45 (80/176), respectively, for clinically significant cancer (Gleason score 3 + 4) and 0.64 (209/329) and 0.47 (42/89), respectively, for cancer (all Gleason scores).

When using MRI Likert scores (especially scores 4 and 5) to detect cancer (Table 3), the sensitivity increased significantly as signal-intensity homogeneity scores improved (reader 1, from 0.412 to 0.714, $p < 0.001$; reader 2, from 0.526 to 0.728, $p = 0.007$), whereas changes in specificity were not statistically significant (reader 1, $p = 0.238$; reader 2, $p = 0.822$). In the detection of significant cancer (Gleason score $\geq 3+4$), the sensitivity also increased with higher (i.e., better) signal-intensity homogeneity scores (reader 1, from 0.5 to 0.824, $p = 0.005$; reader 2, from 0.627 to 0.857, $p = 0.028$), though the specificity decreased significantly for reader 1 ($p = 0.009$).

Discussion

With prostate MRI becoming the method of choice for detection and staging of prostate cancer in most clinical settings and thus being used in an increasing number of patients, background signal intensity changes, seen as hypointense patchy or striated areas on T2-weighted images, are a common sight on routine prostate MRI examinations. These changes occasionally mimic cancerous lesions, but their presence can also obfuscate tumors and hinder their detection [4, 7]. Though the assessment and categorization of factors that may impair tumor detection is part of the standard reporting process for imaging examinations of other cancers (for example, breast tissue density grades are routinely included in screening mammography reports), background signal intensity changes in prostatic tissue are not assessed in a standardized manner [8].

In this study, we sought to develop a standardized classification system to quantify the amount of changes present on prostate MRI and to assess their influence on cancer detection rates. We found significantly lower sensitivity in the detection of both any cancer and clinically significant cancer (Gleason score $\geq 3 + 4$) in patients with a large amount of changes in their peripheral zones and thus poor signal-intensity homogeneity scores (score 2). Conversely, sensitivity reached levels greater than 80% for the detection of significant cancers in patients with minimal or no changes (score ≥ 4). These results held for both readers' assessments of signal-intensity homogeneity scores, with substantial agreement seen.

We also saw a significant decrease in specificity in the detection of significant cancer with better signal-intensity homogeneity scores for one reader ($p = 0.009$) and the same trend for the other reader, though it did not reach statistical significance ($p = 0.127$). A possible explanation is that radiologists may be more confident calling small lesions (even those that prove to be benign on histopathology) suspicious in patients with few or no background changes, thus potentially overcalling cancer in these patients.

The negative effect of postbiopsy hemorrhage on the sensitivity of prostate MRI has been reported [5, 9–11]. However, although the presence of hemorrhagic changes could be mitigated by extending the time between biopsy and MRI (though the exact amount of time necessary is still debated), signal-intensity homogeneity changes on T2-weighted images can be seen in both biopsy-naïve patients and in patients who have undergone biopsy a substantial period before MRI. By categorizing prostate signal-intensity homogeneity, the reporting radiologist informs the ordering clinician of a factor that could limit the diagnostic

sensitivity of the scan in a particular patient. As a result, the scan could be repeated (i.e., in cases where prostatitis is suspected) or at least considered with caution in the clinical context. A scan with abundant signal-intensity homogeneity changes might also prompt supplemental screening, which is already being performed for patients with high breast tissue density as part of breast cancer screening [12]. Although the Prostate Imaging Reporting and Data System version 2 (PI-RADSv2) scoring system indirectly mentions some of our criteria, for example, by grading “linear, wedge-shaped” areas as unlikely to be cancer (PI-RADSv2 score 2), the current system does not entail adjusting the score on the basis of the overall presence of changes in signal-intensity homogeneity [13].

Our findings could have an impact on clinical studies investigating the accuracy of MRI in cancer detection; prior studies incorporate MRI examinations with potentially differing amounts of signal-intensity homogeneity changes into a single study cohort. However, an uneven composition of the cohort with scans of differing amounts of background changes might skew the results, and the sensitivity and specificity of MRI in cancer detection could prove to be significantly different when results are stratified by the amount of changes present. Therefore, integration of a scoring system into structured prostate MRI reports could be of great value to clinical research as well as routine clinical care, facilitating informed clinical decision-making, particularly in patients undergoing active surveillance or when MRI is used for targeting biopsies. In turn, the use of such a scoring system could improve clinician confidence in the modality, thus further promoting appropriate use of prostate MRI.

Our study had limitations. The purpose of the study was to derive a standardized scoring system for background signal intensity changes that could readily be used in routine clinical practice. As such, we did not include assessment of diffusion-weighted or dynamic contrast-enhanced MRI (though those sequences were available to the originally reporting radiologist), and we used a qualitative scoring system. Although T2-weighted sequences are important for localization of suspicious lesions, the addition of scoring for other sequence types could further improve the scoring system, albeit while complicating its use. Still, our results indicate that the proposed scoring system based on T2-weighted images alone is already clinically useful, though validation of the system in a different cohort is desirable. In addition, we limited our assessment to changes of the peripheral zone, because potential changes in the transition zone can be difficult to differentiate from entities such as stromal benign prostate hyperplasia and because the majority of prostate cancers are detected in the peripheral zone.

In conclusion, our study illustrates the importance of prostate signal-intensity homogeneity on T2-weighted images on prostate cancer detection rates and provides a standardized scoring system that can be readily used in routine clinical practice. The use of this scoring system could aid in the clinical management of patients with suspected prostate cancer as well as in the design of studies investigating the performance of prostate MRI in cancer detection.

Acknowledgment

We thank Ada Muellner for editing the manuscript.

Supported in part by National Institutes of Health/National Cancer Institute Cancer Center Support Grant P30 CA008748.

References

1. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *N Engl J Med* 2018; 378:1767–1777 [PubMed: 29552975]
2. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet* 2017; 389:815–822 [PubMed: 28110982]
3. Valerio M, Donaldson I, Emberton M, et al. Detection of clinically significant prostate cancer using magnetic resonance imaging-ultrasound fusion targeted biopsy: a systematic review. *Eur Urol* 2015; 68:8–19 [PubMed: 25454618]
4. Li Y, Mongan J, Behr SC, et al. Beyond prostate adenocarcinoma: expanding the differential diagnosis in prostate pathologic conditions. *RadioGraphics* 2016; 36:1055–1075 [PubMed: 27315446]
5. White S, Hricak H, Forstner R, et al. Prostate cancer: effect of postbiopsy hemorrhage on interpretation of MR images. *Radiology* 1995; 195:385–390 [PubMed: 7724756]
6. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003; 138:168–175 [PubMed: 12558355]
7. Sfanos KS, Yegnasubramanian S, Nelson WG, de Marzo AM. The inflammatory microenvironment and microbiome in prostate cancer development. *Nat Rev Urol* 2018; 15:11–24 [PubMed: 29089606]
8. Freer PE. Mammographic breast density: impact on breast cancer risk and implications for screening. *RadioGraphics* 2015; 35:302–315 [PubMed: 25763718]
9. Park KK, Lee SH, Lim BJ, Kim JH, Chung BH. The effects of the period between biopsy and diffusion-weighted magnetic resonance imaging on cancer staging in localized prostate cancer. *BJU Int* 2010; 106:1148–1151 [PubMed: 20346052]
10. Rosenkrantz AB, Mussi TC, Hindman N, et al. Impact of delay after biopsy and post-biopsy haemorrhage on prostate cancer tumour detection using multi-parametric MRI: a multi-reader study. *Clin Radiol* 2012; 67:e83–e90 [PubMed: 22981729]
11. Tamada T, Sone T, Jo Y, et al. Prostate cancer: relationships between postbiopsy hemorrhage and tumor detectability at MR diagnosis. *Radiology* 2008; 248:531–539 [PubMed: 18539890]
12. Berg WA, Zhang Z, Lehrer D, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 2012; 307:1394–1404 [PubMed: 22474203]
13. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging Reporting and Data System 2015, version 2. *Eur Urol* 2016; 69:16–40 [PubMed: 26427566]

Fig. 1—.
Flowchart detailing patient inclusion process. TRUS = transrectal ultrasound.

Fig. 2—. Examples of images with different prostate signal-intensity homogeneity scores. Changes are evaluated on T2-weighted images and scored on scale of 1–5. Scores of 1 and 2 indicate significant amount of changes in peripheral zone, and score of 5 indicates homogeneous T2 hyperintense peripheral zone.

A, Example of prostate signal-intensity homogeneity score of 1. Images with this score have very poor homogeneity, with diffuse low-signal-intensity changes and markedly limited assessment of focal prostatic lesions.

B, Example of prostate signal-intensity homogeneity score of 2. Images with this score have poor homogeneity, with multifocal or patchy (or both) low-signal-intensity changes and limited assessment of focal prostatic lesions.

C, Example of prostate signal-intensity homogeneity score of 3. Images with this score have moderate homogeneity, with ill-defined or mild (or both) low-signal-intensity changes and mildly limited assessment of focal prostatic lesions.

D, Example of prostate signal-intensity homogeneity score of 4. Images with this score have very good homogeneity, with linear or very small (or both) low-signal-intensity changes and marginally limited assessment of focal prostatic lesions.

E, Example of prostate signal-intensity homogeneity score of 5. Images with this score have excellent homogeneity, with no low-signal-intensity changes and no limitation in assessment of focal prostatic lesions.

Fig. 3—. Additional examples of different prostate signal homogeneity scores.
A, 57-year-old man with prostate signal-intensity homogeneity score of 1.
B, 59-year-old man with prostate signal-intensity homogeneity score of 2.
C, 66-year-old man with prostate signal-intensity homogeneity score of 3.
D, 73-year-old man with prostate signal-intensity homogeneity score of 4.
E, 62-year-old man with prostate signal-intensity homogeneity score of 5.

TABLE 1:
Patient Characteristics, Biopsy Outcomes, and Prostate Signal-Intensity Homogeneity Scores

Reader, Characteristic	Prostate Signal-Intensity Homogeneity Score			P
	2	3	4	
Reader 1				
No. of patients	96	177	145	
Patient age at time of biopsy (y)	62 (41–76)	63 (45–79)	64 (45–77)	0.107
Prebiopsy PSA value (ng/mL)	5 (0–36)	5 (1–28)	6 (0–22)	0.141
Time between MRI and biopsy (d)	47 (0–319)	43 (0–523)	51 (0–408)	0.944
Time between PSA and biopsy (d)	91 (4–364)	91 (1–496)	71 (1–511)	0.965
No. of biopsy cores	16 (4–21)	16 (0–21)	16 (4–20)	0.286
No. of positive biopsy cores	2 (0–19)	2 (0–12)	2 (0–13)	0.234
Percentage of tumor in core	35 (2–96)	30 (2–100)	30 (2–95)	0.153
Length of tumor in core (mm)	4 (0–5)	4 (0–31)	4 (0–32)	0.747
Prebiopsy MRI Likertscore				< 0.001
2	0 (0.0)	3 (1.7)	8 (5.5)	
3	56 (58.3)	53 (29.9)	42 (29.0)	
4	40 (41.7)	111 (62.7)	89 (61.4)	
5	0 (0.0)	10 (5.6)	6 (4.1)	
Biopsy Gleason score				0.212
6	42 (52.5)	73 (53.3)	61 (54.5)	
7	35 (43.8)	61 (44.5)	47 (42.0)	
8	3 (3.8)	1 (0.7)	2 (1.8)	
9	0 (0.0)	2 (1.5)	2 (1.8)	
Patients with Gleason score 3+4	38 (47.5)	64 (46.7)	51 (45.5)	0.316
Patients with prostate cancer	80 (83.3)	137 (77.4)	112 (77.2)	0.514
Reader2				
No. of patients	140	158	120	
Patient age at time of biopsy (y)	63 (41–77)	62 (48–78)	64 (45–79)	0.180
Prebiopsy PSA value (ng/mL)	5 (0–36)	5 (0–21)	6 (0–22)	0.306
Time between MRI and biopsy (d)	40 (0–319)	55 (0–523)	40 (0–408)	0.555
Time between PSA and biopsy (d)	92 (1–364)	86 (1–496)	72 (3–511)	0.577

Reader, Characteristic	Prostate Signal-Intensity Homogeneity Score			P
	2	3	4	
No. of biopsy cores	16 (3–21)	16 (0–20)	16 (4–20)	0.121
No. of positive biopsy cores	2 (0–19)	2 (0–12)	2 (0–13)	0.168
Percentage of tumor in core	38 (2–98)	30 (2–100)	29 (2–95)	0.052
Length of tumor in core (mm)	5 (0–25)	4 (0–31)	4 (0–32)	0.476
Prebiopsy MRI Likertscore				< 0.001
2	0 (0.0)	3 (1.9)	8 (6.7)	
3	67 (47.9)	53 (33.5)	31 (25.8)	
4	68 (48.6)	96 (60.8)	76 (63.3)	
5	5 (3.6)	6 (3.8)	5 (4.2)	
Biopsy Gleason score				0.295
6	57 (49.1)	69 (57.0)	50 (54.3)	
7	55 (47.4)	48 (39.7)	40 (43.5)	
8	4 (3.4)	1 (0.8)	1 (1.1)	
9	0 (0.0)	3 (2.5)	1 (1.1)	
Patients with Gleason score 3+4	59 (50.9)	52 (43.0)	42 (45.7)	0.219
Patients with prostate cancer	116 (82.9)	121 (76.6)	92 (76.7)	0.201

Note—Numeric values are given as medians with ranges in parentheses or absolute values with percentages in parentheses. Information on cores is provided for the most aggressive lesion detected. PSA = prostate-specific antigen.

TABLE 2:

Interreader Agreement on Prostate Signal-Intensity Homogeneity Scores

Score Assigned by Reader 1	Score Assigned by Reader 2					Total
	1	2	3	4	5	
1	7 (2)	6 (1)	0 (0)	1 (0)	0 (0)	14 (3)
2	11 (3)	62 (15)	9 (2)	0 (0)	0 (0)	82 (20)
3	1 (0)	53 (13)	114 (27)	9 (2)	0 (0)	177 (42)
4	0 (0)	0 (0)	35 (8)	90 (22)	5 (1)	130 (31)
5	0 (0)	0 (0)	0 (0)	8 (2)	7 (2)	15 (4)
Total	19 (5)	121 (29)	158 (38)	108 (26)	12 (3)	418

Note—Numbers in parentheses are percentages. The estimated weighted kappa was 0.783 (95% CI, 0.733–0.818) between the two readers.

TABLE 3:
Sensitivity, Specificity, PPV, and NPV of Prebiopsy MRI Likert Scores in Prostate Signal-Intensity Homogeneity Score Subgroups

Cancer Type, Score	Sensitivity	P	Specificity	P	PPV	P	NPV	P
Reader 1								
Significant cancer		0.005		0.009		0.412		0.160
PSHS 2	0.50 (19/38)		0.67 (28/42)		0.58 (19/33)		0.60 (28/47)	
PSHS3	0.81 (52/64)		0.40 (29/73)		0.54 (52/96)		0.71 (29/41)	
PSHS 4	0.82 (42/51)		0.38 (23/61)		0.53 (42/80)		0.72 (23/32)	
Cancer		< 0.001		0.238		0.807		0.080
PSHS 2	0.41 (33/80)		0.56 (9/16)		0.83 (33/40)		0.16 (9/56)	
PSHS3	0.70 (96/137)		0.38 (15/40)		0.79 (96/121)		0.27 (15/56)	
PSHS 4	0.71 (80/112)		0.55 (18/33)		0.84 (80/95)		0.36 (18/50)	
Reader2								
Significant cancer		0.028		0.127		0.073		0.110
PSHS 2	0.63 (37/59)		0.58 (33/57)		0.61 (37/61)		0.60 (33/55)	
PSHS3	0.77 (40/52)		0.41 (28/69)		0.49 (40/81)		0.70 (28/40)	
PSHS 4	0.86 (36/42)		0.38 (19/50)		0.54 (36/67)		0.76 (19/25)	
Cancer		0.007		0.822		0.860		0.143
PSHS 2	0.53 (61/116)		0.50 (12/24)		0.84 (61/73)		0.18 (12/67)	
PSHS3	0.67 (81/121)		0.43 (16/37)		0.79 (81/102)		0.29 (16/56)	
PSHS 4	0.73 (67/92)		0.50 (14/28)		0.83 (67/81)		0.36 (14/39)	

Note—PPV = positive predictive value, NPV = negative predictive value, PSHS = prostate signal-intensity homogeneity score.