

Assessing Extraprostatic Extension with Multiparametric MRI of the Prostate: Mehralivand Extraprostatic Extension Grade or Extraprostatic Extension Likert Scale?

Lars A. R. Reisaeter, MD • Ole J. Halvorsen, MD, PhD • Christian Beisland, MD, PhD • Alfred Honoré, MD • Karsten Gravdal, MD, PhD • Are Losnegård, PhD • Jan Monssen, BSc • Lars A. Akslen, MD, PhD • Martin Biermann, MD, PhD

From the Departments of Radiology (L.A.R.R., A.L., J.M., M.B.), Pathology (O.J.H., K.G., L.A.A.), and Urology (C.B., A.H.), Haukeland University Hospital, Bergen, Norway; Department of Clinical Medicine (L.A.R.R., C.B., A.L., M.B.) and Centre for Cancer Biomarkers CCBIO (O.J.H., L.A.A.), University of Bergen, Jonas Liesvei 65, N-5021 Bergen, Norway. Received August 19, 2019; revision requested September 27; final revision received October 4; accepted October 21. **Address correspondence to** L.A.R.R. (e-mail: lars.reisaeter@helse-bergen.no).

L.A.R.R. supported by the Western Norway Regional Health Authority (Helse Vest RHF).

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

See also the commentary by Choyke in this issue.

Radiology: Imaging Cancer 2020; 2(1):e190071 • <https://doi.org/10.1148/rycan.2019190071> • Content codes: **GU** **MR** **OI**

Purpose: To validate the MRI grading system proposed by Mehralivand et al in 2019 (the “extraprostatic extension [EPE] grade”) in an independent cohort and to compare the Mehralivand EPE grading system with EPE interpretation on the basis of a five-point Likert score (“EPE Likert”).

Materials and Methods: A total of 310 consecutive patients underwent multiparametric MRI according to a standardized institutional protocol before radical prostatectomy was performed by using the same 1.5-T MRI unit at a single institution between 2010 and 2012. Two radiologists blinded to clinical information assessed EPE according to standardized criteria. On the basis of the readings performed until 2017, the diagnostic performance of EPE Likert and Mehralivand EPE score were compared using receiver operating characteristics (ROC) and decision curve methodology against histologic EPE as standard of reference. Prediction of biochemical recurrence-free survival (BRFS) was assessed by Kaplan-Meier analysis and log rank test.

Results: Of the 310 patients, 80 patients (26%) had EPE, including 33 with radial distance 1.1 mm or greater. Interrater reliability was fair (weighted κ 0.47 and 0.45) for both EPE grade and EPE Likert. Sensitivity for identifying EPE using EPE grade versus EPE Likert was 0.83 versus 0.86 and 0.86 versus 0.91 for radiologist 1 and 2, respectively. Specificity was 0.48 versus 0.58 and 0.39 versus 0.70 ($P < .05$ for radiologist 2). There were no significant differences in the ROC area under the curve or on decision curve analysis. Both EPE grade and EPE Likert were significant predictors of BRFS.

Conclusion: Mehralivand EPE grade and EPE Likert have equivalent diagnostic performance for predicting EPE and BRFS with a similar degree of observer dependence.

Supplemental material is available for this article.

© RSNA, 2020

Prostate cancer is the most common noncutaneous malignancy among men in Northern America and Europe (1,2). For intermediate- and high-risk disease, the recommended treatment is external beam radiation therapy or radical prostatectomy (3,4).

At the time of presentation, most patients have organ-confined prostate cancer (OCP) (1). Patients with OCP can benefit from nerve-sparing surgical procedures that will reduce the chance of incontinence and impotence without increasing the risk of positive surgical margins (5,6). It is therefore important to distinguish patients with OCP from patients with non-organ-confined prostate cancer, which includes patients with extraprostatic extension (EPE) (7).

Not only is the presence or absence of EPE an independent predictor of biochemical recurrence-free survival (BRFS), but BRFS is strongly correlated to the histologic severity of EPE (8–14). Epstein et al and Wheeler et al

subdivided EPE into established versus focal EPE (8,9). Later, Sung et al and Danneman et al identified radial distance of EPE as the most robust histopathologic EPE measure to predict BRFS (10,12).

In a preoperative setting, multiparametric MRI can be used to predict EPE with a pooled sensitivity and specificity of 0.57 and 0.91, respectively (15), having superior diagnostic accuracy over standard MRI (16–18) and EPE nomograms based on clinical and biopsy information (19–21).

Traditionally, EPE is measured on a Likert score based on subjective overall assessment of a combination of multiparametric MRI criteria by an experienced radiologist (22). In the Prostate Imaging Reporting and Data System (PI-RADS) version 2.0 guidelines, the criteria for EPE are (a) abutment, (b) irregularity and neurovascular bundle thickening, (c) bulge, loss of capsule, and capsular enhancement, (d) measurable extracapsular disease, and (e) obliteration of

Abbreviations

ADC = apparent diffusion coefficient, AUC = area under the curve, BRFS = biochemical recurrence-free survival, EPE = extraprostatic extension, IQR = interquartile range, PI-RADS = Prostate Imaging Reporting and Data System, RALP = robotic-assisted laparoscopic prostatectomy, ROC = receiver operating characteristics

Summary

The Mehralivand grading system has reliable performance applied in a different cohort and performs just as well as subjective extraprostatic extension Likert scores; both systems rely on experience-dependent criteria (eg, bulging and irregularity), thus subjectivity is not ruled out.

Key Points

- Mehralivand extraprostatic extension grading is a valid grading system of extraprostatic extension and has similar diagnostic performance as a five-point Likert assessment of extraprostatic extension.
- Using both systems for assessing extraprostatic extension, higher scores reflected a higher likelihood of extraprostatic extension, but did not correlate well with extraprostatic extension severity expressed as radial distance of extraprostatic extension in the histopathologic specimen.
- A large prospective multicentric trial of multiparametric MRI criteria for extraprostatic extension is needed under the auspices of a major urologic society.

the rectoprostatic angle (23). Concerning measurable EPE, both curvilinear contact length (24–26) and apparent diffusion coefficient (ADC) (27–30) have been identified as reliable predictors of EPE. The relative weighting of these criteria, however, is not explicit and confined to the realm of expert opinion.

In a recent study, Mehralivand et al proposed a new grading system for EPE (“EPE grade”), based on a set of potentially less observer-dependent criteria: curvilinear contact length, irregularity, and bulging. Against histologic findings as a standard of reference, the system achieved 75% sensitivity and 68% specificity for EPE grade 1 or greater in a retrospective cohort. The Mehralivand EPE grading system is based on a measurement of curvilinear contact length and the binary EPE criteria of bulging, irregularity, and “visible” EPE. EPE grade 1 is defined as either curvilinear contact length of 1.5 cm or greater or capsular bulge or irregularity. EPE grade 2 is defined as curvilinear contact length of 1.5 cm or greater in combination with capsular bulge or irregularity. Finally, EPE grade 3 is defined as visible EPE at MRI (31).

The aim of our study was to evaluate the diagnostic performance of Mehralivand EPE grade in an independent prospective cohort from another institution against histopathology and BRFS as reference standards and compare it against Likert scoring of EPE by two radiologists.

Materials and Methods

Patient Population

Of 362 consecutive patients with prostate cancer who underwent robotic-assisted laparoscopic radical prostatectomy (RALP) between January 1, 2010, and December 31, 2012, 310 patients had undergone preoperative multiparametric

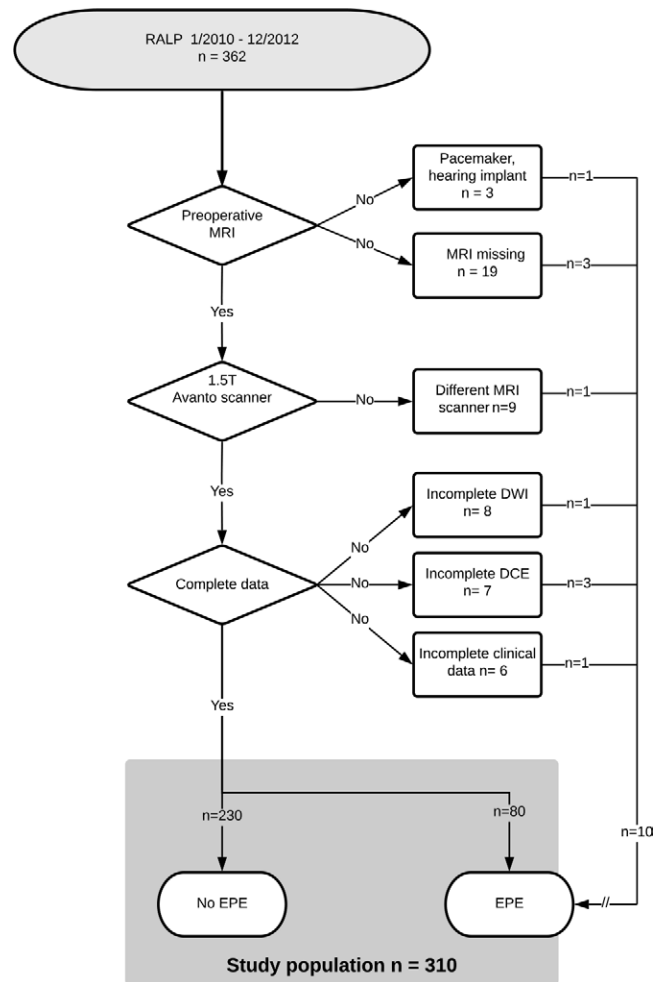


Figure 1: Standards for Reporting of Diagnostic Accuracy (or STARD) flow-chart of inclusion and exclusion. DCE = dynamic contrast-enhanced images, DWI = diffusion-weighted images, EPE = extraprostatic extension, RALP = robotic-assisted laparoscopic prostatectomy.

MRI using the same 1.5-T MR machine by a standardized protocol read by the same two dedicated radiologists, and had complete clinical and follow-up data for analysis (Fig 1). A subset of 63 patients who underwent surgery in 2010 from the same cohort has been published previously regarding localization of the index tumor in preoperative multiparametric MRI using PI-RADS version 1 (32). The present cohort was included in a larger study of 591 patients that examined multiparametric MRI criteria for optimizing risk stratification after RALP with BRFS but not EPE as endpoint (33). The study had been approved by the institutional review board, and all patients had given informed written consent.

Imaging Technique

All patients were examined using the same 1.5-T MR machine (Avanto; Siemens Medical Systems, Erlangen, Germany), using an integrated endorectal and pelvic phased-array coil (MR Innerva; Medrad, Pittsburgh, Pa) for signal reception. The scanning protocol included T2-weighted series, diffusion-weighted images, and dynamic contrast material–enhanced images (Table 1).

Table 1: Imaging Protocol for 1.5-T MRI with an Endorectal Coil

Sequence	Plane	Repetition Time/ Echo Time (msec)	Intersection Gap (mm)	Matrix	Field of View (mm)	Acquisition Time (min)
T2 weighted	Sagittal	3030/98	0.8	320 × 256	200 × 200	3:06
T2 weighted	Coronal	3000/98	0.4	320 × 256	200 × 200	4:05
T2 weighted	Axial	4840/84	0.8	320 × 256	200 × 200	4:18
VIBE	Axial	7.23/2.55	0.8	192 × 192	250 × 250	0:20
DWI ($b = 50, 400, 800 \text{ sec/mm}^2$)	Axial	3000/72	0.8	128 × 128	128 × 128	5:33
DWI ($b = 1200 \text{ sec/mm}^2$)	Axial	2800/83	0.6	128 × 128	250 × 250	2:23
DCE-TWIST + C	Axial	4.24/1.66	0.8	512 × 512	192 × 138	6:58

Note.—DCE-TWIST + C = dynamic contrast-enhanced, time-resolved interleaved stochastic trajectories sequence with a time resolution of 6.16 seconds using intravenous contrast material, DWI = diffusion-weighted imaging, VIBE = volumetric interpolated breath-hold examination.

MRI Evaluation and Reporting

All multiparametric MRI examinations were read in random order in single contiguous sessions by each of two dedicated radiologists (J. Rørvik and L.A.R.R.), who were blinded to each other and the result of the histopathologic examination of the surgical specimens, in a time period between 2011 and 2015 for a previously published study (33). The EPE criteria—asymmetry of the neurovascular bundle, capsular irregularity, and capsular bulging—were assessed in the course of the same reading session with separate scores for each criterion by each radiologist using a five-point Likert score (1 = criterion not present, 2 = probably not present, 3 = uncertain if present, 4 = probably present, 5 = definitely present) including an overall score. The overall EPE score was based on the overall subjective assessment of presence of EPE by each of the observers at the time of reading the multiparametric MRI, using a five-point Likert score. In addition, both radiologists measured tumor curvilinear contact length and tumor size on axial T2-weighted images using a standard PACS system with diagnostic screens (Agfa Impax v.6.5; Agfa Health Care, Mortsel, Belgium) and recorded the PI-RADS version 1 scores of the leading lesions.

Histopathologic Evaluation

Two uropathologists (O.J.H. and K.G.) independently outlined the presence of tumor involvement on drawings of whole-mount step sections of the entire prostate taken at 5-mm intervals. For each specimen, the pathologists determined the presence of a pathologic index tumor, and volume of the tumors was estimated using routine pathologic measurements, as previously described (32).

In addition, the pathologists determined the number of sections with EPE, location, maximum radial distance, and circumferential length. Each instance of EPE in a given specimen was subclassified into focal and nonfocal EPE (34). Extensive EPE was defined as EPE with radial extension 1.1 mm or greater based on Danneman et al (12).

Histopathologic findings of the preoperative biopsies were recorded in the study database, as previously published (32,33). Following the 2014 International Society of Urological Pathology consensus meeting, Gleason grades and scores were aggregated into five grade groups (35).

Data Storage

A custom-developed MDCake database was used to collect all data since 2010 using dedicated data entry forms with drop-down lists for all categorical data (32,33,36). All data used were collected in the database before October 2017 and then frozen.

Generation of Mehralivand EPE Grades

Likert scores for the separate EPE criteria (bulging, irregularity, and asymmetry of the neurovascular bundle) and curvilinear contact length, which had been recorded individually by the two blinded observers, were transformed into Mehralivand EPE grades 0–3 independently for each patient and each observer in April 2019 according to Mehralivand criteria published in January 2019 (31) by a Structured Query Language script at the level of the relational database.

Statistical Analysis

All statistics were calculated in R (37) using packages {epiR}, {irr}, {pROC}, {survival}, {MASS}, and {rmda} as previously published (32,33,36). Data were summarized using descriptive statistics by tabulating the median, means, and interquartile ranges (IQRs) for continuous variables. To assess the interrater reliability of multiparametric MRI criteria, we used weighted κ values. To assess agreement of the curvilinear contact-length, we used Bland-Altman analysis (38). The diagnostic performance of multiparametric MRI was evaluated by calculating sensitivity, specificity, positive predictive value, and negative predictive value, with confidence intervals estimated by bootstrapping with 2000 samples. Receiver operating characteristics (ROC) were used to visualize diagnostic performance, and the DeLong test was used to check for

significant differences in the area under the curve (AUC) between EPE grade and EPE Likert.

Furthermore, EPE grade and EPE Likert were tabulated against the quartiles of histopathologic radial distance. Ideally, readings should be clustered along the diagonal of the diagram. For qualitative assessment, false-positive outliers were defined as the three cells in the lower left corner with high scores and limited radial distance and false-negative outliers as the three cells in the upper right corner with low scores but extensive radial distance.

Differences in the Kaplan-Meier plots for BRFS were tested using log rank after dichotomizing EPE grade at the level of 1 and EPE Likert at the level of 3. Finally, decision curve analysis was applied to assess differences between dichotomized EPE grade and EPE Likert with and without clinical features. The significance level for all statistical tests was 5% (two-sided).

Results

Patient Characteristics and Endpoints

Of the 310 patients who met the inclusion criteria, EPE was present in 80 patients (26%). Among these, 33 had extensive EPE with radial distance of 1.1 mm or greater. Median histologic index tumor size was 2.1 mL (IQR, 0.8–4.2 mL). The remaining patient characteristics are summarized in Table 2.

EPE was localized mostly dorsolaterally and particularly at the base (Fig E1 [supplement]). Surgical margins were positive in 46 of 310 patients (15%), 27 in category pT3 (27 of 80; 34%) and 19 in pT2 (19 of 230; 8%). Thirteen patients had serum prostate-specific antigen level 0.02 ng/mL or greater after RALP. Of the 297 patients with initial biochemical remission, 23 had biochemical recurrence: 11 of 230 (5%) without EPE, 12 of 80 (15%) with EPE, including eight of 33 (24%) with extensive EPE ($P < .01$). Histopathologic extensive EPE was a significant predictor of BRFS (Fig E2 [supplement]).

Interrater Agreement

Interrater agreement between radiologists 1 and 2 was weak to moderate with a weighted κ of 0.47 for Mehrlivand EPE grade and 0.45 for EPE Likert. For diagnosing probable EPE (EPE grade ≥ 1 , EPE Likert ≥ 3), Cohen κ was 0.38 and 0.54, respectively. Bland-Altman analysis of measurements of curvilinear contact length showed poor agreement between readers (Fig 2).

Diagnostic Performance

Average sensitivity across both readers was 92.5% for EPE grade versus 77% for EPE Likert with an average specificity of 42.5% versus 64%, respectively. When comparing readings for each radiologist, there was a trend toward increased sensitivity at the cost of decreased specificity of the Mehrlivand EPE grade, which reached statistical significance for only one of the observers, radiologist 2 (Table 3).

Table 2: Characteristics of 310 Patients

Parameter	Value
Preoperative	
Age at biopsy (y)	63.1 (59–66)*
MRI to surgery (d)	24.5 (8–82)*
Percentage of positive biopsy cores	34.5 (20–50)*
cT stage (DRE)	
T1c	209
T2a	49
T2b	33
T2c	9
T2a/b	10
Capra score	
Low	78
Intermediate	187
High	45
ISUP grade group	
1	128
2	127
3	32
4	19
5	4
PI-RADS v1 score leading lesion for radiologist 1/radiologist 2	
>3	19/36
3	13/39
4	40/79
5	238/156
Postoperative	
Age at surgery (y)	63.6 (60–67)*
Preoperation s-PSA (ng/mL)	8.8 (6–13)*
Histologic findings (TNM 2007, seventh edition) pT	
T2a	16
T2b	9
T2c	199
T3a	66
T3b	19 [†]
T4	1 [‡]
ISUP grade group	
1	52
2	180
3	60
4	8
5	10

Note.—Unless otherwise indicated, data are numbers of patients. cT = clinical T stage, DRE = digital rectal examination, EPE = extraprostatic extension, ISUP = International Society of Urological Pathology, PI-RADS v1 = Prostate Imaging Reporting and Data System, version 1, pT = pathologic T stage, sPSA = serum prostate-specific antigen. * Data are medians, with interquartile ranges in parentheses. [†] Four patients with seminal vesicle invasion did not have EPE. [‡] Macroscopic invasion of the bladder.

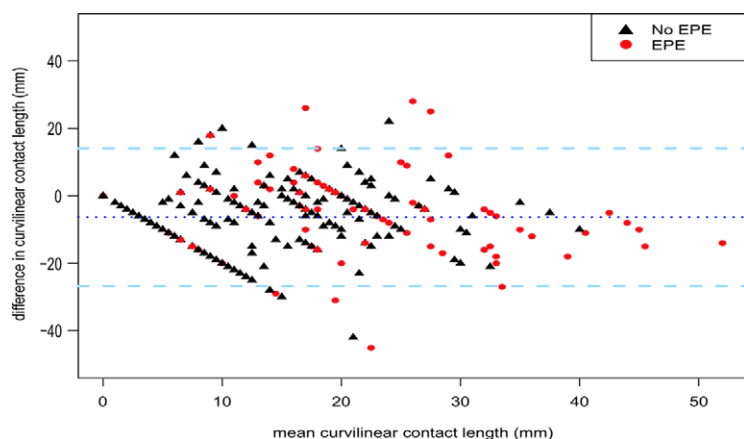


Figure 2: Bland-Altman plot of measurement of curvilinear contact length by two radiologists. The mean of the two measurements in millimeters (x-axis) is plotted against difference between the two measurements (y-axis). Mean difference of the two measurements and limits of agreement (± 2 standard deviations) are indicated by the stippled dark and light blue lines, respectively. Red circles = measurements in patients with histopathologic extraprostatic extension (EPE), black triangles = measurements in patients without EPE.

Table 3: Diagnostic Performance of Mehralivand EPE Grade and EPE Likert

Parameter and Radiologist	All EPE				EPE with Radial Distance ≥ 1.1 mm*			
	EPE Grade ≥ 1		EPE Likert ≥ 3		EPE Grade ≥ 1		EPE Likert ≥ 3	
	Value	95% CI	Value	95% CI	Value	95% CI	Value	95% CI
Sensitivity								
Radiologist 1	0.86	0.77, 0.93	0.82	0.72, 0.90	0.97	0.84, 1.00	0.94	0.80, 0.99
Radiologist 2	0.88	0.78, 0.94	0.72	0.61, 0.81	0.97	0.84, 1.00	0.91	0.76, 0.98
Specificity								
Radiologist 1	0.47	0.41, 0.54	0.58	0.52, 0.65	0.43	0.37, 0.49	0.53	0.47, 0.59
Radiologist 2	0.38	0.32, 0.45	0.70	0.63, 0.75	0.35	0.29, 0.41	0.65	0.59, 0.70
PPV								
Radiologist 1	0.36	0.30, 0.44	0.41	0.33, 0.49	0.17	0.12, 0.23	0.19	0.13, 0.26
Radiologist 2	0.33	0.27, 0.40	0.45	0.36, 0.54	0.15	0.11, 0.21	0.23	0.16, 0.32
NPV								
Radiologist 1	0.91	0.84, 0.95	0.91	0.85, 0.95	0.99	0.96, 1.00	0.99	0.95, 1.00
Radiologist 2	0.90	0.82, 0.95	0.88	0.82, 0.92	0.99	0.95, 1.00	0.98	0.95, 1.00

Note.—CI = confidence interval, EPE = extraprostatic extension, NPV = negative predictive value, PPV = positive predictive value.

* Danneman et al (12).

AUC for detecting histopathologic EPE using EPE grade was 0.80 and 0.75 (not significant) and EPE Likert 0.76 and 0.77 (not significant) for radiologist 1 and 2, respectively (Fig 3). Likewise, there were no significant differences between AUC for predicting extensive EPE (Fig E3 [supplement]).

Clinical Outcomes

Against BRFS as a secondary endpoint, EPE grade and EPE Likert performed similarly well for each of the radiologists (not significant; Fig 4).

Qualitative Assessment

When tabulating individual EPE grade and EPE Likert scores against quartiles of radial distance of EPE in an evaluation matrix, outliers occurred for both EPE assessment methods and both observers with frequencies ranging between 20% and 38% (Fig 5). Although radiologist 1 had a tendency toward higher sensitivity, radiologist 2 tended

toward higher specificity at the expense of sensitivity with either EPE assessment method. Two representative cases, one true-positive and one false-negative, are shown in Figures 6 and 7.

Discussion

The present study provided validation of the new grading system proposed by Mehralivand et al for predicting histopathologic EPE at preoperative multiparametric MRI in an independent cohort. We used histopathologic findings and BRFS as primary and secondary reference standards. In addition, we compared Mehralivand EPE grade with previously recorded EPE Likert assessments by two blinded radiologists.

To permit standardized reading of multiparametric MRI with regard to malignancy, PI-RADS has been jointly released by the American College of Radiology and the European Society of Urogenital Radiology. PI-RADS has fair diagnostic performance (40) with moderate reproducibility for experienced

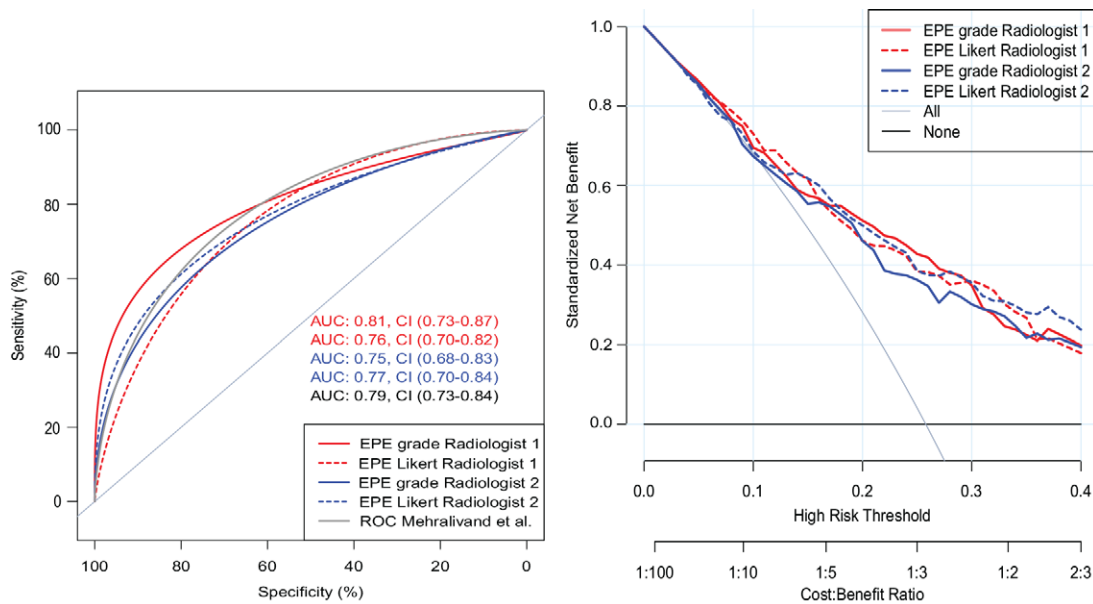


Figure 3: (a) Receiver operating characteristic (ROC) analysis and (b) decision curve analysis of Mehralivand extraprostatic extension (EPE) grade compared with EPE Likert for each of two observers with histopathologic EPE as reference standard. (a) For convenience, the original Mehralivand ROC curve is included in light gray. (b) Combined clinical score University of California San Francisco Cancer of the Prostate Risk Assessment (UCSF-CAPRA) (39) and EPE grade/EPE Likert. AUC = area under the curve.

radiologists (41). However, PI-RADS version 2.0 does not yet include an explicit scoring system for EPE assessment (42). The new Mehralivand EPE grading system could provide a template.

As Mehralivand et al point out, previous “conventional” assessment of EPE relied on an overall five-point Likert assessment of a set of imaging criteria for EPE by an experienced radiologist (23). By necessity, such an assessment was subjective because there were no explicit weighting factors for the different criteria. In this context, it is interesting to note that the Likert score originated as a research tool in cognitive psychology (43).

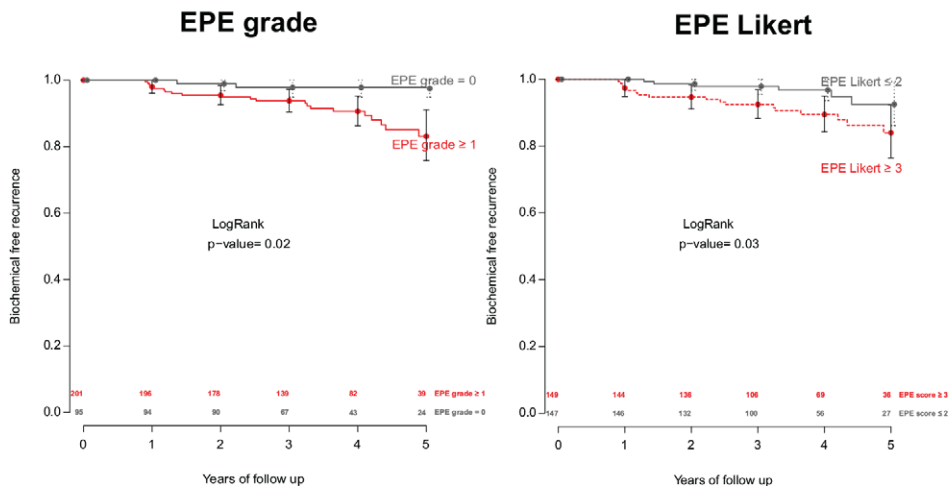
The Mehralivand EPE grading system achieved similar sensitivity and specificity in our cohort as in the original publication, and no difference in the ROC curves was observed (Fig 3). In comparison with our previously recorded EPE Likert scores, the performance of the Mehralivand system was equivalent, but not superior. Although frank breach of prostate capsule or invasion into periprostatic space or adjacent anatomic structures represent straightforward criteria most radiologists will agree on, both systems are based on the more subtle assessment of bulging and irregularity. Thus, subjectivity is not ruled out with either approach. Surprisingly, even quantitative measurements were fraught with interobserver error. Bland-Altman analysis showed that our measurements of curvilinear contact length—which are crucial in respect to the Mehralivand EPE grading system—differed between the two radiologists, at times markedly so. The reasons for the discrepancies were difficult to ascertain retrospectively because we did not document the precise location of the measurements taken; the largest discrepancies can probably be attributed to measuring different index lesions.

Ideally, curvilinear contact length should have been recorded for each tumor. Qualitative analysis in the form of an evaluation matrix revealed that the EPE grade was burdened with as much subjective bias between the two radiologists as the EPE Likert. Independent of the system used, radiologist 1 tended toward higher sensitivity and radiologist 2 toward higher specificity. Both systems had equivalent performance for predicting BRFS.

It was unclear how the Mehralivand threshold of curvilinear contact length at 15 mm was derived. Other similar studies proposed thresholds of curvilinear capsule length that are in the range of 12–20 mm, based on the median in the observed distribution (24–26). Other quantitative measures, such as tumor size and ADC, have been identified as predictors of EPE (27–30). In addition, Lim et al and Alessi et al have shown that the PI-RADS lesion score can predict or rule out EPE (44,45).

What should an ideal scoring system incorporate? Criteria and measurements should be based on precise definitions, easy to apply in a clinical setting, robust, and with a high level of interrater agreement. Investigation should include looking at both quantitative measures (abutment, ADC, tumor-to-gland volume ratio) as well as the qualitative criteria from PI-RADS version 2. Thresholds should be carefully calibrated in relation to scanner systems and institutional cohort achieving an optimal balance between sensitivity and specificity based on decision curve analysis. Intuitively, higher scores should not only indicate a higher likelihood of EPE but be correlated with higher severity of EPE. Also suggested is adopting radial extension of EPE in histopathologic step sections rather than a binary criterion as the reference standard. This would facilitate multivariate multilevel regression models that can

Radiologist 1



Radiologist 2

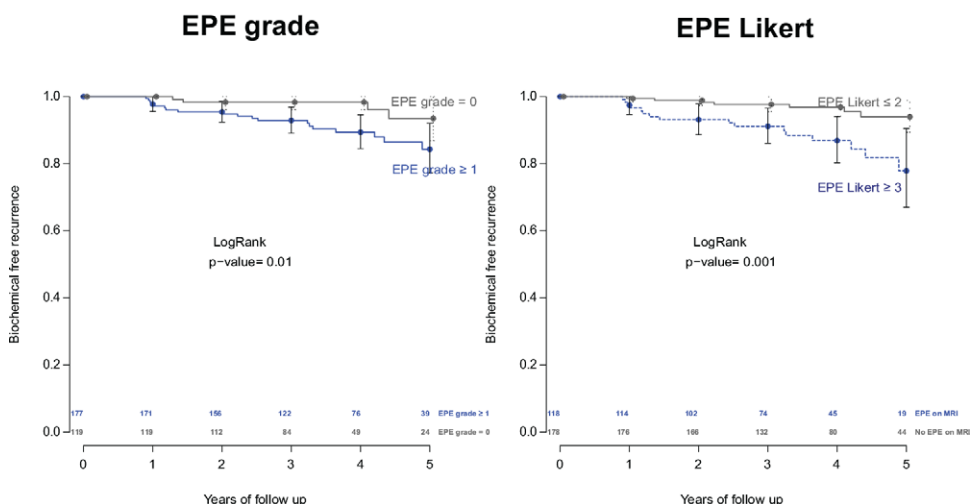


Figure 4: Kaplan-Meier plots of biochemical recurrence-free survival with extraprostatic extension (EPE) grade 1 or greater and EPE Likert score 3 or greater for each of the radiologists. Differences were not significant.

then serve as templates for building an optimized EPE grading system.

The strengths of our evaluation were that it was based on a homogeneous cohort, prospectively collected over just 3 consecutive years at a single institution. The study adhered to a strictly enforced standardized protocol with a relatively small number of exclusions. EPE assessment by the pathologists was standardized not only regarding the presence of EPE but also regarding the radial distance of EPE. BRFS was included as a secondary independent reference standard. Reading of images, including measurement of curvilinear contact length by each observer, followed a consistent standard and was blinded in relation to clinical data, outcomes, and future classification systems. Encoding of the readings was performed using a dedicated database application with a high degree of granularity ahead of time, permitting a prospective and lossless transformation of the original readings into the proposed new Mehralivand scoring system.

Limitations of the present study included the following: (a) We did not attempt to launch yet another grading system given the limitations of single-cohort data. This would call for a large prospective multicenter trial under the auspices of a major urology society. (b) Because of the death of radiologist (J. Rørvik) in 2018, the cohort was limited to 310 patients, and could not be expanded without introducing bias. (c) Treatment allocation of patients was in accordance with the then-current guidelines (46). (d) All patients were scanned on the same 1.5-T system using an endorectal coil, in contrast to Mehralivand et al using 3-T with an endorectal coil. We have since moved over to 3-T with dedicated pelvic coils. (e) Transforming previous historic readings into a different scoring system, such as the Mehralivand system, has the disadvantage that the readings under the two compared systems were not independent. However, the strength of this approach was that it isolated the differences between the competing scoring systems, eliminating the confounders intrarater variability and

	Radiologist 1				Radiologist 2			
	(0-0.5)	[0.5-1.2)	[1.2-1.7)	[1.7-∞)	(0-0.5)	[0.5-1.2)	[1.2-1.7)	[1.7-∞)
EPE grade 3	12	14	18	17	0	0	4	4
EPE grade 2	0	0	0	0	3	2	1	4
EPE grade 1	5	1	0	2	12	14	14	12
EPE grade 0	3	5	2	1	5	4	1	0
Outliers	26	+	5	=39%	3	+	13	=20%
EPE Likert 5	8	11	11	14	0	0	1	4
EPE Likert 4	5	2	7	4	2	1	4	5
EPE Likert 3	2	2	0	0	8	9	10	10
EPE Likert 1 & 2	5	5	2	2	10	10	5	1
Outliers	24	+	4	=35%	2	+	16	=23%

Figure 5: Evaluation matrix assessing extraprostatic extension (EPE) grade and EPE Likert, based on quartiles of radial distance (n = 80). Radial distance (mm), (0-0.5) = first quartile, (0.5-1.2) = second quartile, (1.2-1.7) = third quartile, (1.7-∞) = fourth quartile. Light gray = outliers with high score/grade and a small radial distance. Dark gray = outliers with low score/grade and a large radial distance.

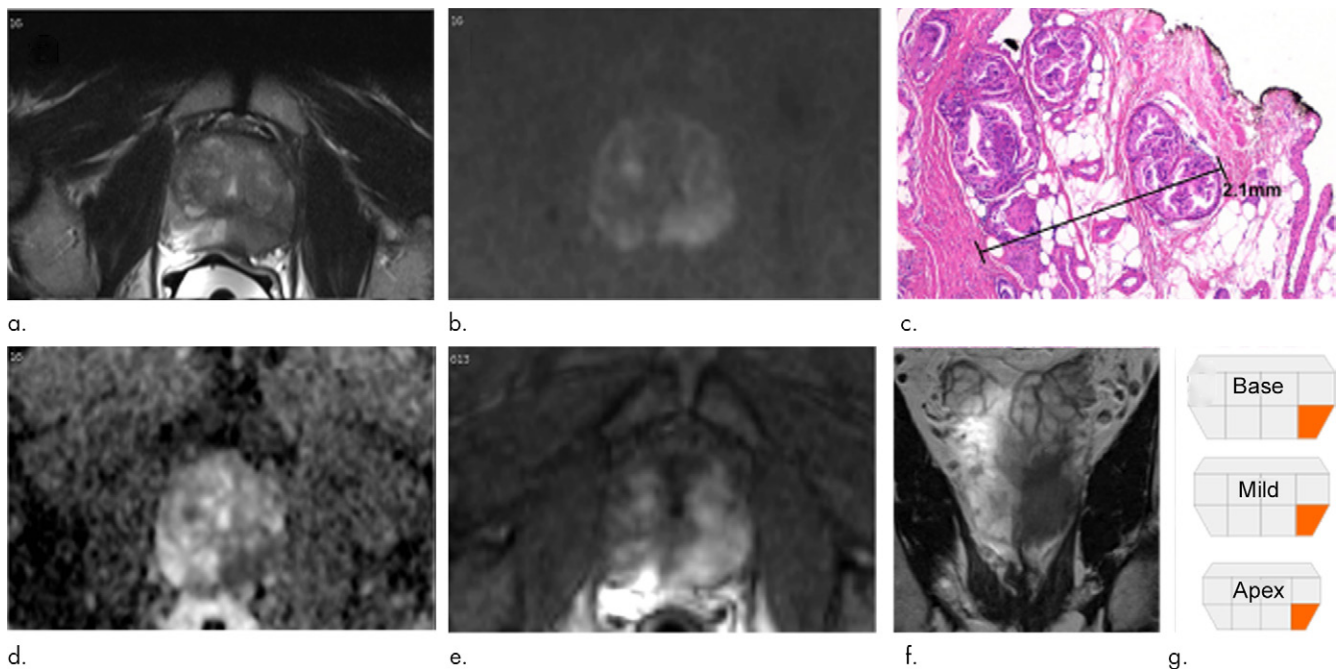


Figure 6: Multiparametric MR and histopathologic images in a 62-year-old patient with extensive extraprostatic extension (EPE). (a) T2-weighted MR image, (b) image with calculated $b = 1200 \text{ sec/mm}^2$, (c) hematoxylin-eosin-stained step-section image showing EPE with a radial distance of 2.1 mm, (d) apparent diffusion coefficient map, (e) dynamic contrast-enhanced image, (f) coronal T2-weighted image, and (g) histopathologic image showing localization of EPE (orange). Multiparametric MRI EPE grade was 3 and EPE Likert score was 5 according to both radiologists.

recall bias that would have occurred if we had reread the examinations in two separate sessions using different scoring systems.

In conclusion, the present study validated the recently proposed Mehrlivand EPE grading system against an independent institutional cohort using histopathologic findings

and BRFS as primary and secondary reference standard. Comparison with previously recorded EPE assessment on a five-point Likert score showed equivalent diagnostic performance, however, with similar degrees of observer dependence.

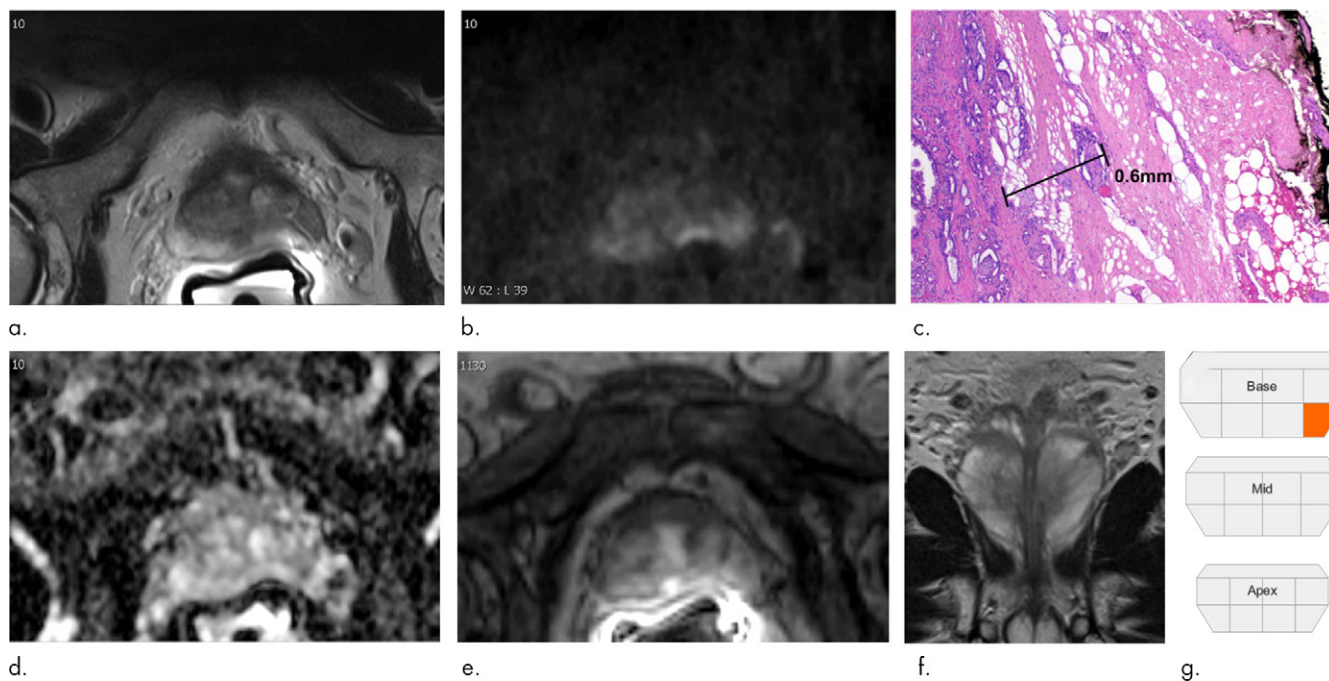


Figure 7: Multiparametric MR and histopathologic images in a 67-year-old patient with limited extraprostatic extension (EPE) with a radial distance of 0.6 mm overlooked by both radiologists. EPE grade was 0 and EPE Likert score was 0. Conventions are as in Figure 6.

Acknowledgment: We would like to thank the late Prof. J. Rørvik, who read MRI examinations for this study.

Author contributions: Guarantors of integrity of entire study, L.A.R.R., M.B.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, L.A.R.R., O.J.H., C.B., K.G., J.M., L.A.A.; clinical studies, L.A.R.R., C.B., A.H., J.M., L.A.A.; statistical analysis, L.A.R.R., C.B., M.B.; and manuscript editing, all authors.

Disclosures of Conflicts of Interest: L.A.R.R. disclosed no relevant relationships. O.J.H. disclosed no relevant relationships. C.B. disclosed no relevant relationships. A.H. disclosed no relevant relationships. K.G. disclosed no relevant relationships. A.L. disclosed no relevant relationships. J.M. disclosed no relevant relationships. L.A.A. disclosed no relevant relationships. M.B. disclosed no relevant relationships.

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68(1):7–30.
- Ferlay J, Colombet M, Soerjomataram I, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer* 2018;103:356–387.
- Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol* 2017;71(4):618–629.
- Sanda MG, Cadeddu JA, Kirkby E, et al. Clinically localized prostate cancer: AUA/ASTRO/SUO guideline. Part I: risk stratification, shared decision making, and care options. *J Urol* 2018;199(3):683–690.
- Harris CR, Punnen S, Carroll PR. Men with low preoperative sexual function may benefit from nerve sparing radical prostatectomy. *J Urol* 2013;190(3):981–986.
- Noldus J, Michl U, Graefen M, Haese A, Hammerer P, Huland H. Patient-reported sexual function after nerve-sparing radical retropubic prostatectomy. *Eur Urol* 2002;42(2):118–124.
- Kattan MW, Wheeler TM, Scardino PT. Postoperative nomogram for disease recurrence after radical prostatectomy for prostate cancer. *J Clin Oncol* 1999;17(5):1499–1507.
- Epstein JI, Carmichael MJ, Pizov G, Walsh PC. Influence of capsular penetration on progression following radical prostatectomy: a study of 196 cases with long-term followup. *J Urol* 1993;150(1):135–141.
- Wheeler TM, Dillioglulugil O, Kattan MW, et al. Clinical and pathological significance of the level and extent of capsular invasion in clinical stage T1-2 prostate cancer. *Hum Pathol* 1998;29(8):856–862.
- Sung MT, Lin H, Koch MO, Davidson DD, Cheng L. Radial distance of extraprostatic extension measured by ocular micrometer is an independent predictor of prostate-specific antigen recurrence: a new proposal for the substaging of pT3a prostate cancer. *Am J Surg Pathol* 2007;31(2):311–318.
- Chan SM, Garcia FJ, Chin JL, Moussa M, Gabril MY. The clinical significance of in-depth pathological assessment of extraprostatic extension and margin status in radical prostatectomies for prostate cancer. *Prostate Cancer Prostatic Dis* 2011;14(4):307–312.
- Danneman D, Wiklund F, Wiklund NP, Egevad L. Prognostic significance of histopathological features of extraprostatic extension of prostate cancer. *Histopathology* 2013;63(4):580–589.
- Ball MW, Partin AW, Epstein JI. Extent of extraprostatic extension independently influences biochemical recurrence-free survival: evidence for further pT3 subclassification. *Urology* 2015;85(1):161–164.
- Jeong BC, Chalfin HJ, Lee SB, et al. The relationship between the extent of extraprostatic extension and survival following radical prostatectomy. *Eur Urol* 2015;67(2):342–346.
- de Rooij M, Hamoen EHJ, Witjes JA, Barentsz JO, Rovers MM. Accuracy of magnetic resonance imaging for local staging of prostate cancer: a Diagnostic Meta-analysis. *Eur Urol* 2016;70(2):233–245.
- Bloch BN, Furman-Haran E, Helbich TH, et al. Prostate cancer: accurate determination of extracapsular extension with high-spatial-resolution dynamic contrast-enhanced and T2-weighted MR imaging—initial results. *Radiology* 2007;245(1):176–185.
- Boesen L, Chabanova E, Løgager V, Balslev I, Mikines K, Thomsen HS. Prostate cancer staging with extracapsular extension risk scoring using multiparametric MRI: a correlation with histopathology. *Eur Radiol* 2015;25(6):1776–1785.
- Bianchi R, Cozzi G, Petralia G, et al. Multiparametric magnetic resonance imaging and frozen-section analysis efficiently predict upgrading, upstaging, and extraprostatic extension in patients undergoing nerve-sparing robotic-assisted radical prostatectomy. *Medicine (Baltimore)* 2016;95(40):e4519.

19. Wang L, Hricak H, Kattan MW, Chen HN, Scardino PT, Kuroiwa K. Prediction of organ-confined prostate cancer: incremental value of MR imaging and MR spectroscopic imaging to staging nomograms. *Radiology* 2006;238(2):597–603.
20. Chen Y, Yu W, Fan Y, et al. Development and comparison of a Chinese nomogram adding multi-parametric MRI information for predicting extracapsular extension of prostate cancer. *Oncotarget* 2017;8(13):22095–22103.
21. Morlacco A, Sharma V, Viers BR, et al. The incremental role of magnetic resonance imaging for prostate cancer staging before radical prostatectomy. *Eur Urol* 2017;71(5):701–704.
22. Hricak H, Wang L, Wei DC, et al. The role of preoperative endorectal magnetic resonance imaging in the decision regarding whether to preserve or resect neurovascular bundles during radical retropubic prostatectomy. *Cancer* 2004;100(12):2655–2663.
23. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* 2016;69(1):16–40.
24. Baco E, Rud E, Vlatkovic L, et al. Predictive value of magnetic resonance imaging determined tumor contact length for extracapsular extension of prostate cancer. *J Urol* 2015;193(2):466–472.
25. Rosenkrantz AB, Shanbhogue AK, Wang A, Kong MX, Babb JS, Taneja SS. Length of capsular contact for diagnosing extraprostatic extension on prostate MRI: assessment at an optimal threshold. *J Magn Reson Imaging* 2016;43(4):990–997.
26. Kongnyuy M, Sidana A, George AK, et al. Tumor contact with prostate capsule on magnetic resonance imaging: a potential biomarker for staging and prognosis. *Urol Oncol* 2017;35(1):30.e1–30.e8.
27. Lawrence EM, Gallagher FA, Barrett T, et al. Preoperative 3-T diffusion-weighted MRI for the qualitative and quantitative assessment of extracapsular extension in patients with intermediate- or high-risk prostate cancer. *AJR Am J Roentgenol* 2014;203(3):W280–W286.
28. Kim CK, Park SY, Park JJ, Park BK. Diffusion-weighted MRI as a predictor of extracapsular extension in prostate cancer. *AJR Am J Roentgenol* 2014;202(3):W270–W276.
29. Lim C, Flood TA, Hakim SW, et al. Evaluation of apparent diffusion coefficient and MR volumetry as independent associative factors for extraprostatic extension (EPE) in prostatic carcinoma. *J Magn Reson Imaging* 2016;43(3):726–736.
30. Krishna S, Lim CS, McInnes MDF, et al. Evaluation of MRI for diagnosis of extraprostatic extension in prostate cancer. *J Magn Reson Imaging* 2018;47(1):176–185.
31. Mehralivand S, Shih JH, Harmon S, et al. A grading system for the assessment of risk of extraprostatic extension of prostate cancer at multiparametric MRI. *Radiology* 2019;290(3):709–719.
32. Reisäter LA, Fütterer JJ, Halvorsen OJ, et al. 1.5-T multiparametric MRI using PI-RADS: a region by region analysis to localize the index-tumor of prostate cancer in patients undergoing prostatectomy. *Acta Radiol Stockh Swed* 1987 2015;56(4):500–511.
33. Reisäter LAR, Fütterer JJ, Losnegård A, et al. Optimising preoperative risk stratification tools for prostate cancer using mpMRI. *Eur Radiol* 2018;28(3):1016–1026.
34. Epstein JI, Partin AW, Sauvageot J, Walsh PC. Prediction of progression following radical prostatectomy. A multivariate analysis of 721 men with long-term follow-up. *Am J Surg Pathol* 1996;20(3):286–292.
35. Epstein JI, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol* 2016;40(2):244–252.
36. Biermann M. A simple versatile solution for collecting multidimensional clinical data based on the CakePHP web application framework. *Comput Methods Programs Biomed* 2014;114:70–79.
37. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2015. <http://www.R-project.org/>. Accessed August 1, 2014.
38. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307–310.
39. Cooperberg MR, Pasta DJ, Elkin EP, et al. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. *J Urol* 2005;173(6):1938–1942.
40. Zhang L, Tang M, Chen S, Lei X, Zhang X, Huan Y. A meta-analysis of use of Prostate Imaging Reporting and Data System Version 2 (PI-RADS V2) with multiparametric MR imaging for the detection of prostate cancer. *Eur Radiol* 2017;27(12):5204–5214.
41. Rosenkrantz AB, Ginocchio LA, Cornfeld D, et al. Interobserver reproducibility of the PI-RADS Version 2 Lexicon: a multicenter study of six experienced prostate radiologists. *Radiology* 2016;280(3):793–804.
42. Turkbey B, Rosenkrantz AB, Haider MA, et al. Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* 2019;76(3):340–351.
43. Likert R. A Technique for the Measurement of Attitudes. *Arch Psychol* 1932;140:1–55.
44. Lim CS, McInnes MDF, Lim RS, et al. Prognostic value of Prostate Imaging and Data Reporting System (PI-RADS) v. 2 assessment categories 4 and 5 compared to histopathological outcomes after radical prostatectomy. *J Magn Reson Imaging* 2017;46(1):257–266.
45. Alessi S, Pricolo P, Summers P, et al. Low PI-RADS assessment category excludes extraprostatic extension (³pT3a) of prostate cancer: a histology-validated study including 301 operated patients. *Eur Radiol* 2019;29(10):5478–5487.
46. Heidenreich A, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. *Eur Urol* 2011;59(1):61–71.