Korean Journal of Radiology

The Diagnostic Performance of the Length of Tumor Capsular Contact on MRI for Detecting Prostate Cancer Extraprostatic Extension: A Systematic Review and Meta-Analysis

Tae-Hyung Kim, MD[1, 2], Sungmin Woo, MD[3], Sangwon Han, MD[4]*, Chong Hyun Suh, MD[5], Soleen Ghafoor, MD[3], Hedvig Hricak, MD[3], Hebert Alberto Vargas, MD[3]

[1]Department of Radiology, Seoul National University College of Medicine, Seoul, Korea; [2]Department of Radiology, Naval Pohang Hospital, Pohang, Korea; [3]Department of Radiology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA; [4]Department of Nuclear Medicine, Armed Forces Daejeon Hospital, Daejeon, Korea; [5]Department of Radiology and Research Institute of Radiology, Asan Medical Center, Seoul, Korea

Objective: The purpose was to review the diagnostic performance of the length of tumor capsular contact (LCC) on magnetic resonance imaging (MRI) for detecting prostate cancer extraprostatic extension (EPE).
Materials and Methods: PubMed and EMBASE databases were searched up to March 24, 2019. We included diagnostic accuracy studies that evaluated LCC on MRI for EPE detection using radical prostatectomy specimen histopathology as the reference standard. Quality of studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool. Sensitivity and specificity were pooled and graphically presented using hierarchical summary receiver operating characteristic (HSROC) plots. Meta-regression and subgroup analyses were conducted to explore heterogeneity.
Results: Thirteen articles with 2136 patients were included. Study quality was generally good. Summary sensitivity and specificity were 0.79 (95% confidence interval [CI] 0.73–0.83) and 0.67 (95% CI 0.60–0.74), respectively. Area under the HSROC was 0.81 (95% CI 0.77–0.84). Substantial heterogeneity was present among the included studies according to Cochran's Q-test ($p < 0.01$) and Higgins I^2 (62% and 86% for sensitivity and specificity, respectively). In terms of heterogeneity, measurement method (curvilinear vs. linear), prevalence of Gleason score $\geq$ 7, MRI readers' experience, and endorectal coils were significant factors ($p \leq 0.01$), whereas method to determine the LCC threshold, cutoff value, magnet strength, and publication year were not ($p = 0.14$–0.93). Diagnostic test accuracy estimates were comparable across all assessed MRI sequences.
Conclusion: Greater LCC on MRI is associated with a higher probability of prostate cancer EPE. Due to heterogeneity among the studies, further investigation is needed to establish the optimal cutoff value for each clinical setting.
Keywords: *Prostate cancer; Magnetic resonance imaging; Meta-analysis; Systematic review*

INTRODUCTION

Assessment of extraprostatic extension (EPE) in prostate cancer patients is associated with adverse oncological outcomes such as post-treatment biochemical recurrence, development of metastasis, and decreased survival (1-3). Recognition of the presence of EPE is critical in patients treated with radical prostatectomy (RP) (e.g., whether to

perform nerve-sparing procedures) and treatment planning for patients who undergo radiotherapy (4). However, it is challenging to accurately predict EPE based only on clinical assessment using digital rectal examination, biopsy Gleason scores, and/or prostate-specific antigen levels (5, 6).

Multiparametric magnetic resonance imaging (mp-MRI) has been widely utilized for the detection, local staging, and treatment planning in patients with prostate cancer (7). However, the accuracy of mp-MRI in determining EPE has been variable among studies (8). This may stem from the fact that EPE evaluation on MRI has traditionally been based on the subjective assessment of imaging findings of abutment, irregularity or prostate capsule, bulging, and neurovascular bundle thickening on T2-weighted images (T2WIs). Poor interreader agreement and dependence on the level of experience renders consistent reporting of EPE among radiologists difficult (9, 10). To overcome this shortcoming of subjective EPE evaluation, objective and quantitative measures have been introduced for predicting EPE on mp-MRI, including calculation of apparent diffusion coefficient (ADC) values of the dominant lesion or standardizing interpretation and reporting using the Prostate Imaging Reporting and Data System (PI-RADS) (11, 12).

The length of tumor capsular contact (LCC), defined as the length of prostate tumor in contact with the capsule, has been proposed as an independent and reproducible predictor of EPE (13-16). LCC showed improved accuracy in predicting EPE compared to the previously reported qualitative MRI findings of bulging or irregularity of the capsule (13). However, its adoption in clinical practice has been slow for several reasons. Prior studies have been based on small numbers of patients and have assessed LCC using different thresholds (6–20 mm) using various MRI sequences (i.e., T2WI, ADC, or dynamic contrast-enhanced [DCE] MRI) (13-16). The purpose of this study was to systematically review the literature and perform a meta-analysis regarding the diagnostic performance of LCC on MRI for detecting EPE in prostate cancer.

MATERIALS AND METHODS

This study was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis of Diagnostic Test Accuracy guidelines (17). A research question based on the patient, index test, comparator, outcome, and study design (PICOS) criteria was formulated as follows: what is the diagnostic performance of LCC on

MRI for predicting EPE in prostate cancer patients, as compared with that of histopathological results after RP, in original articles?

Literature Search
PubMed and EMBASE databases were systematically searched up to March 24, 2019. The following keywords and related terms were included in the search query: prostat* AND ("magnetic resonance" OR MR OR MRI) AND (capsul* OR contact) AND (extracapsular OR extraprostatic). The references of identified articles were screened to find other eligible studies.

Inclusion Criteria
Studies were considered eligible in the meta-analysis if they met the following PICOS criteria (18): 1) patients diagnosed with prostate cancer; 2) LCC on MRI used for EPE detection as an index test; 3) histopathological results after RP as comparator, 4) EPE as outcome; and 5) original articles as type of study.

Exclusion Criteria
Exclusion criteria were 1) studies with fewer than ten patients; 2) publication type other than original articles; 3) studies focusing on different topics (i.e., diagnostic accuracy of other MRI findings for EPE prediction); 4) overlapping patient populations; and 5) insufficient data necessary for meta-analytic pooling (even after attempts to contact the authors). If overlap was present among multiple publications, the study with the largest patient cohort was included. Two reviewers performed the literature search and study selection independently. Consensus was reached after discussion with a third reviewer.

Data Extraction and Quality Assessment
Data regarding patient, study, and MRI characteristics were extracted using a standardized form. Methodological quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool (18). Two reviewers independently performed both data extraction and quality assessment followed by discussion with a third reviewer in cases of disagreement.

Data Synthesis and Analysis
Sensitivity and specificity were calculated from reconstructed data from the included studies in 2 x 2 tables (true positive, false negative, false positive, and

true negative). Results from the more experienced reader were used for the meta-analysis if results from multiple independent readers were available.

The pooled sensitivity and specificity and their 95% confidence interval (CI) were calculated using the bivariate random effects model (19). A hierarchical summary receiver operating characteristics (HSROCs) curve with a 95% confidence region and prediction region was presented graphically to display the results (20). Publication bias was assessed using the Deeks' funnel plot and Deeks' asymmetry test (21).

Heterogeneity was determined using both Cochran's Q test with $p < 0.05$ indicating statistical significance (22) and inconsistency index (I^2) using the following criteria (23): 0–40%, heterogeneity might not be important; 30–60% moderate heterogeneity may be present; 50–90%, substantial heterogeneity may be present; and 75–100%, considerable heterogeneity. The threshold effect was visually assessed with coupled forest plots of sensitivity and specificity and a Spearman correlation coefficient greater than 0.6 between sensitivity and false-positive rates was considered to suggest a considerable threshold effect (24).

Meta-regression analysis was conducted to explore the cause of heterogeneity using the following categories: 1) method to determine threshold of LCC (receiver operating characteristics [ROCs] curve vs. others); 2) LCC cutoff value (> 10 mm vs. ≤ 10 mm based on 10 mm in PI-RADS version 2 and > 12 vs. ≤ 12 mm dichotomized by the median of

included studies); 3) LCC measurement method (curvilinear vs. linear); 4) prevalence of high Gleason score (≥ 7) on biopsy (≥ 75.4% [median of included studies] vs. < 75.4%); 5) magnetic field strength (3- vs. 1.5-Tesla [T]); 6) experience of MR readers (≥ 10 years vs. < 10 years); 7) use of endorectal coils; and 8) publication year (before 2000 vs. after 2000). Additional subgroup analyses stratified according to the MRI sequences that were used to measure LCC (T2WI, ADC, and DCE) were performed.

The "midas" module in Stata 10.0 (StataCorp LLC, College Station, TX, USA) and "mada" package in R software version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) were used for statistical analyses with $p < 0.05$ suggesting statistical significance.

RESULTS

Literature Search

Initially, 176 studies were found in the systematic literature search. After removing 57 duplicates, screening of the 119 titles and abstracts yielded 30 potentially eligible studies. After full-text reviews, 17 studies were excluded for the following reasons: insufficient data to reconstruct 2 x 2 tables (n = 3), focusing on overall local staging of prostate cancer (n = 7), and assessment of other MRI findings as a predictor of EPE (n = 7). Ultimately, 13 original articles including 2136 patients assessing the diagnostic performance of LCC on MRI for detection of EPE in prostate
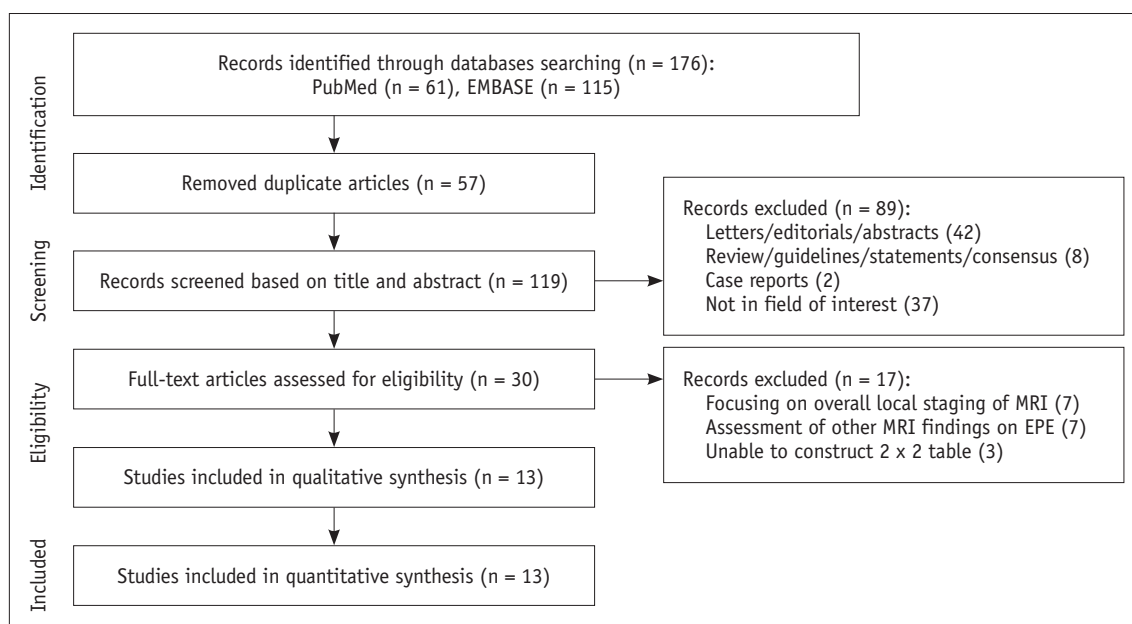

Fig. 1. Flow diagram describing study selection process for meta-analysis. EPE = extraprostatic extension

cancer patients were analyzed (13-16, 25-33). Figure 1 summarizes the detailed study selection process.

Characteristics of Included Studies

The patient characteristics are shown in Table 1. The number of patients ranged from 30 to 553 patients. Seven studies reported the Gleason score from biopsy with a median value of 7. Pathological T stage in addition to histopathological EPE status was reported in seven studies.

The study characteristics are described in Table 2. Only one study was prospective in design, but LCC measurement was performed in a retrospective manner. In terms of the methods to determine optimal LCC threshold, a ROC curve was utilized in ten studies, 10 mm as stated in the PI-RADS v2 guideline in one study, and the method was unclear in two studies. All studies evaluated LCC on T2WI. LCC was additionally measured on ADC in three studies, on DCE in two studies, and the maximum value among all sequences in one study. The threshold value of LCC ranged from 6 to 20 mm. LCC was measured using a curvilinear ruler tool in nine studies, linear in three, and the method was not specified in one.

The MRI characteristics are summarized in Table 3. Eight studies used 3T scanners and five studies used 1.5T scanners. The experience of MRI readers ranged from 0.5 to 22 years. An endorectal coil was utilized in five studies and was not used in eight.

Quality Assessment

In general, the quality of the studies was considered good, with 10 of the 13 studies satisfying five or more of the seven domains (Fig. 2). In the patient selection domain, two studies had an unclear risk of bias, as they did not describe whether patients were consecutively enrolled (15, 32). Regarding the index test domain, there was a high risk of bias in two studies that did not explicitly mention how the LCC cutoff value was determined (31, 33). There was an unclear risk of bias in three studies as it was not explicit whether MRI was read blinded to clinicopathological information (13, 27, 28). Regarding reference standard domain, there was an unclear risk of bias in eight studies as it was unclear whether pathologists were blinded to the MRI interpretation (13, 15, 25-28, 31, 32). Regarding the flow and timing domain, five studies had an unclear risk of bias as the interval between MRI and surgery was not provided (27-31).

Diagnostic Performance of LCC on MRI for Detection of EPE

The range of sensitivities and specificities of the individual studies were 59% to 91% and 44% to 88%, respectively. The Q-test demonstrated that heterogeneity was present ($p < 0.01$). The Higgins I^2 statistics demonstrated substantial heterogeneity in terms of both the sensitivity ($I^2 = 62\%$) and specificity ($I^2 = 86\%$). A threshold effect was not evident based on the coupled forest plots (Fig. 3) and Spearman correlation coefficient between the sensitivity and false-positive rate (-0.174 [95% CI -0.661–0.417]).

For the 13 studies combined, the summary sensitivity was

Table 1. Patient Characteristics of Included Studies

First Author	No. of Patients	Age (Years) Median (Range)	PSA (ng/mL) Median (Range)	Gleason Score Median (Range)	Pathological T Stage[†] (No. of Patients)
Baco (13)	111	64 (45–75)	8.9 (2.5–44)	7 (6–9)	T2a (8), T2b (2), T2c (61), T3a (37), T3b (3)
Caglic (25)	75	65 (57–67)	8.5 (5.7–10.4)	7 (6–10)	T2 (27), T3a (41), T3b (7)
Costa (26)	80	64* (46–78)	8.0* (1.8–46.3)	7 (6, 7, ≥ 8)	T2 (40), T3 (40)
Granja (27)	92	61*(39–78)	9.3* (NA)	NA	T2 (73), T3 (19)
Kongnyuy (28)	379	60 (38–76)	5.5 (0.1–55.7)	7 (6, 7, ≥ 8)	NA
Krishna (29)	149	63* (NA)	7.8* (0.0–73.0)	NA	NA
Matsuoka (30)	210	67 (50–81)	7.0 (2.9–30.0)	7 (5–10)	NA
Mehralivand (31)	553	60* (38–76)	6.3 (0.2–170.0)	NA	NA
Onay (14)	105	62* (40–77)	8.0* (2.1–46.0)	7 (6–9)	T2a (7), T2b (2), T2c (72), T3a (19), T3b (5)
Outwater (32)	30	NA (NA)	11.3* (3.7–30.0)	NA	NA
Rosenkrantz (15)	90	64* (NA)	9.0* (NA)	NA	NA
Woo (16)	185	67* (45–79)	10.2* (0.5–123.0)	7 (6–9)	T2a (32), T2b (3), T2c (94), T3a (41), T3b (15)
Yu (33)	77	63* (45–90)	10.7* (NA)	NA	T2a (6), T2b (9), T2c (28), T3a (22), T3b (5), T3c (7)

*Mean, [†]Reported pathological T stage in addition to histopathological extraprostatic extension status. NA = not available, PSA = prostate-specific antigen

Table 2. Study Characteristics of Included Studies

	Origin				Design		LCC Measurement Methods			
First Author	Institutions	Country	Period	Publication Year	Prospective	Multicenter	Method to Set Threshold	Measured Sequence	Cutoff Value (mm)[†]	Ruler Tool
Baco (13)	Oslo University Hospital	Norway	Jan 2010– Sep 2013	2015	No	No	ROC	T2WI	20	Curvilinear
Caglic (25)	Norfolk and Norwich University Hospital	England	Sep 2014– Jan 2017	2019	No	No	ROC	T2WI	10.5	Curvilinear
Costa (26)	University of Texas Southwestern Medical Center	USA	Nov 2015– Jul 2016	2018	No	No	PI-RADS v2	T2WI	10	Linear
Granja (27)	Hospital Universitario Fundación Santa Fe de Bogotá	Columbia	Mar 2011– Dec 2013	2017	No	No	ROC	T2WI	17.5	Curvilinear
Kongnyuy (28)	National Institutes of Health, Bethesda	USA	May 2007– Dec 2015	2017	No	No	ROC	T2WI	12.5	Curvilinear
Krishna (29)	Ottawa Hospital	Canada	Nov 2012– May 2015	2018	No	No	ROC	T2WI	11	Curvilinear
Matsuoka (30)	Tokyo Medical and Dental University Graduate School	Japan	Aug 2007– Mar 2015	2017	No	No	ROC	T2WI	10	Curvilinear
Mehralivand (31)	University Medical Center, Mainz	Germany	Jun 2007– Mar 2017	2019	No*	No	NA	T2WI	15	Curvilinear
Onay (14)	Baskent University School of Medicine	Turkey	2012–2017	2019	No	No	ROC	T2WI, ADC, DCE	14, 14[‡] 13[§]	Curvilinear
Outwater (32)	Thomas Jefferson University Hospital & Jefferson Medical College	USA	NA	1994	No	No	ROC	T2WI	12	Linear
Rosenkrantz (15)	NYU Langone Medical Center	USA	NA	2016	No	No	ROC	T2WI, ADC	6, 7[‡]	Linear
Woo (16)	Seoul National University Hospital	Korea	Jan 2012– Dec 2012	2016	No	No	ROC	T2WI, ADC, DCE	14, 13[‡] 12[§]	Curvilinear
Yu (33)	University of California	USA	May 1992– Mar 1995	1997	No	No	NA	T2WI	12	NA

*Prospective designed study, but LCC measurement was performed retrospectively, [†]Otherwise specified, threshold set on T2WI, [‡]Threshold set on ADC, [§]Threshold set on DCE. ADC = apparent diffusion coefficient, DCE = dynamic contrast enhanced, LCC = length of tumor capsular contact, PI-RADS = Prostate Imaging Reporting and Data System, ROC = receiver operating characteristic, T2WI = T2-weighted image

0.79 (95% CI 0.73–0.83) with a specificity of 0.67 (95% CI 0.60–0.74). Summary positive likelihood ratio and negative likelihood ratio were 2.4 (95% CI 1.9–3.0) and 0.31 (95% CI 0.24–0.41), respectively. A large difference between the 95% confidence and prediction regions were noted in the HSROC curve, also suggesting heterogeneity between studies (Fig. 4). The area under the HSROC curve was 0.81 (95% CI 0.77–0.84). The likelihood of publication bias was low according to Deeks' funnel plot, with a p-value of 0.69 for the slope coefficient (Fig. 5).

Heterogeneity Exploration

The results of meta-regression analysis are provided in Table 4. Among the several covariates evaluated, LCC

Table 3. MRI Characteristics of Included Studies

First Author	Magnet Strength (T)	Vendor	Model	No. of Readers	Experience (Years)*	Endorectal Coil
Baco (13)	1.5	Siemens	Avanto	1	5, 8	No
Caglic (25)	3	GE	MR750	1	8	No
Costa (26)	3	Philips	Ingenia, Achiva	5	NA	Yes
Granja (27)	1.5	GE	Signa Excite	1	All > 5	No
Kongnyuy (28)	3	Philips	Achieva	2	8, 16	Yes
Krishna (29)	3	Siemens	Trio Tim	2	11, 15	No
Matsuoka (30)	1.5	Philips	Achieva	2	10, 5	No
Mehralivand (31)	3	Philips	Achieva	2	15, 9	Yes
Onay (14)	3	Siemens	Magnetom, Skyra	2	12, 5	No
Outwater (32)	1.5	GE	NA	3	NA	Yes
Rosenkrantz (15)	3	Siemens	Magnetom Trio, Skyra, Biograph	2	1, 4	No
Woo (16)	3	Siemens, Philips	Verio, Trio, Ingenia	1	22	No
Yu (33)	1.5	GE	Signa	3	3, 1, 0.5	Yes

*Result by reader with highest experience was used for analysis except for Costa et al. (26) which provided overall results from 5 readers.
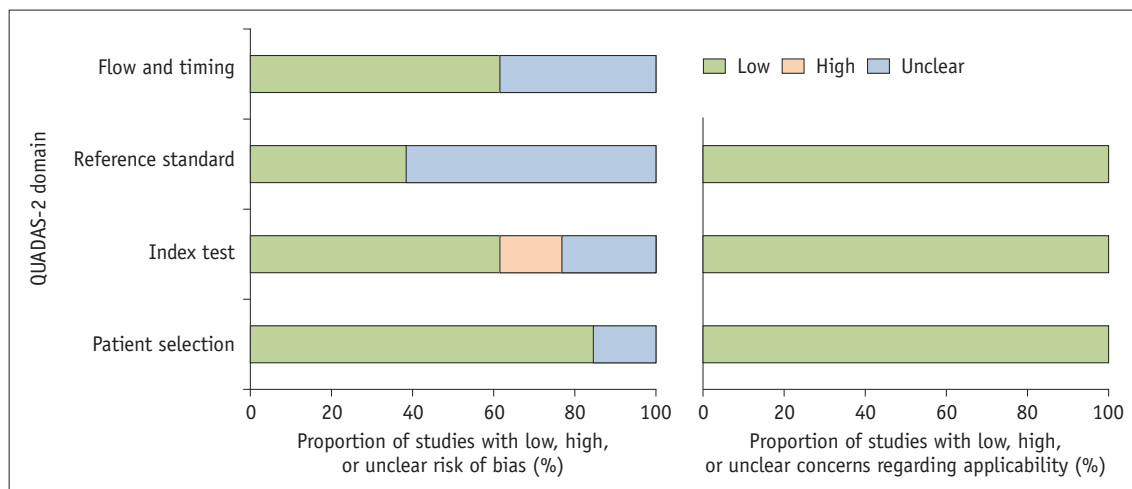

Fig. 2. Grouped bar charts show risk of bias (left) and concerns of applicability (right) for 13 studies using QUADAS-2 tool.
QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies-2

measurement method, prevalence of high Gleason score on biopsy, experience of MR readers, and use of endorectal coils were significant factors affecting heterogeneity ($p \leq$ 0.01). However, when comparing sensitivity and specificity estimates among these subgroups, significant and clinically meaningful differences were only seen regarding endorectal coil usage. Studies that used endorectal coils showed a significantly lower sensitivity (0.72 [95% CI 0.66–0.77]) compared with those not using endorectal coils (0.83 [95% CI 0.79–0.87], $p = 0.001$). For other subgroup comparisons, there were significant overlap in the 95% CIs of the sensitivities and specificities ($p = 0.07–0.83$ and $0.02–0.81$ for sensitivity and specificity, respectively). The method used to determine the LCC threshold, LCC cutoff value, magnet strength, and publication year were not significant

factors affecting the heterogeneity ($p = 0.14–0.93$).

Additional subgroup analyses revealed both sensitivity and specificity estimates were comparable regardless of the MRI sequences used to measure LCC as follows: T2WI (n = 13) (sensitivity 0.78 [95% CI 0.73–0.82], specificity 0.68 [95% CI 0.59–0.75]); ADC (n = 3) (sensitivity 0.77 [95% CI 0.66–0.86], specificity 0.68 [95% CI 0.56–0.77]); DCE (n = 2) (sensitivity 0.80 [95% CI 0.66–0.90], specificity 0.67 [95% CI 0.56–0.77]), and maximum length from all sequences (n = 1) (sensitivity 0.84 [95% CI 0.71–0.93], specificity 0.76 [95% CI 0.68–0.83]).

DISCUSSION

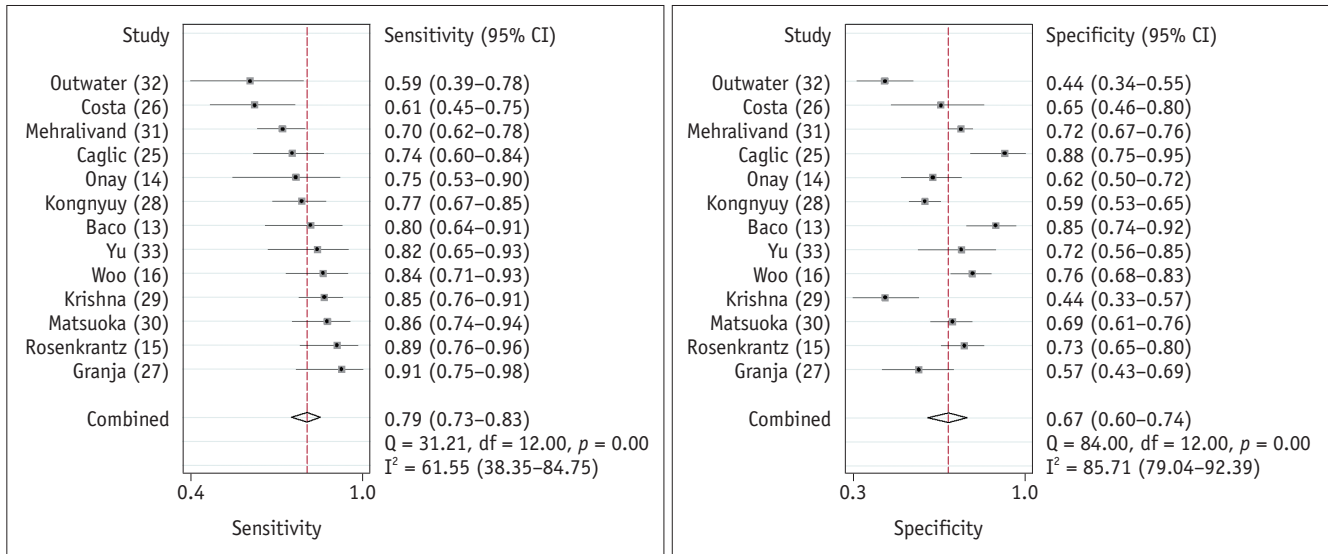In the current study, we performed a meta-analysis on the

Fig. 3. Coupled forest plots of summary sensitivity and specificity. Numbers are pooled estimates with 95% CIs in parentheses. Corresponding heterogeneity statistics are provided in bottom right corners. Horizontal lines indicate 95% CIs. CI = confidence interval
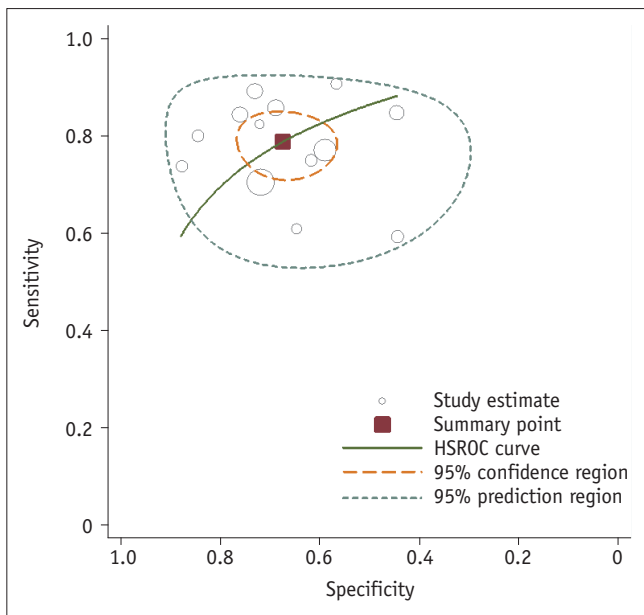

Fig. 4. HSROC curve of diagnostic performance of length of tumor capsular contact on MRI for extracapsular extension prediction in prostate cancer patients. HSROC = hierarchical summary receiver operating characteristic
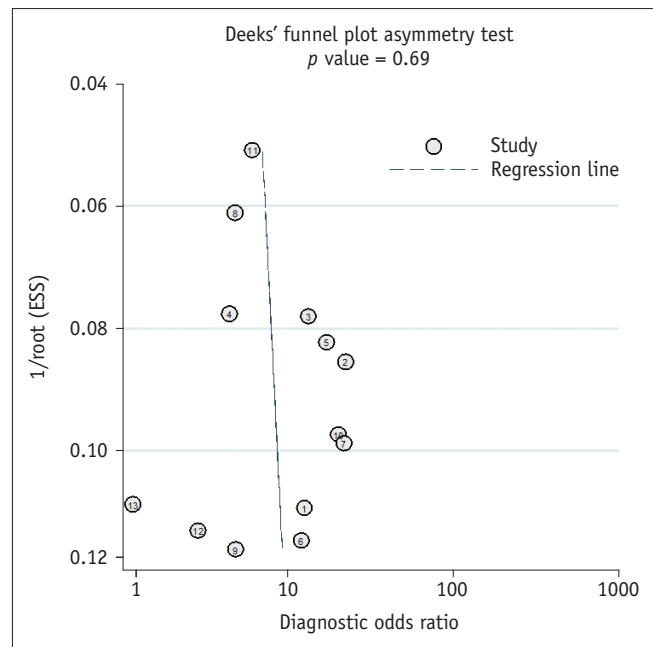

Fig. 5. Deeks' funnel plot. Likelihood of publication bias was low with *p* value of 0.69 for slope coefficient. ESS = effective sample size

diagnostic performance of LCC on MRI for the detection of EPE in prostate cancer. Overall, the diagnostic performance of LCC was moderate with sensitivity and specificity estimates of 0.79 (95% CI 0.73–0.83) and 0.67 (95% CI 0.60–0.74), respectively. Nevertheless, it is noteworthy that the sensitivity was relatively higher than results based on the subjective assessment of EPE (pooled sensitivity of 0.57) as reported in a meta-analysis by de Rooij et al (8). In fact, it has been shown that additionally using LCC along

with other indirect criteria for EPE resulted in increased sensitivity for detecting EPE (57.4% to 83.9%) compared with that when using only direct criteria (i.e., focal capsular irregularity/disruption or neurovascular bundle invasion) (34). This may have important clinical implications as setting either a high sensitivity or specificity reading could be applied to different clinical settings (35). High sensitivity is required when selecting optimal patients to enroll in active surveillance or choosing candidates for RP

Table 4. Meta-Regression Analyses Stratified by Multiple Variables

Variable	No. of Studies	Category	Sensitivity Pooled Value (95% CI)	P	Specificity Pooled Value (95% CI)	P	LRT Chi-Square	P (Joint Model)
Method to set threshold of LCC	10	ROC curve	0.81 (0.77–0.85)	0.073	0.67 (0.59–0.75)	0.685	3.91	0.14
	3	Others*	0.71 (0.62–0.79)		0.70 (0.56–0.84)			
LCC cutoff value	10	> 10 mm	0.78 (0.73–0.84)	0.834	0.67 (0.59–0.75)	0.777	0.14	0.93
	3	≤ 10 mm	0.80 (0.70–0.90)		0.69 (0.55–0.84)			
	6	> 12 mm	0.79 (0.72–0.86)	0.825	0.69 (0.59–0.79)	0.639	0.26	0.88
	7	≤ 12 mm	0.78 (0.72–0.85)		0.66 (0.56–0.76)			
LCC measurement method	9	Curvilinear	0.80 (0.75–0.85)	0.146	0.69 (0.61–0.77)	0.445	12.18	< 0.01
	3	Linear	0.72 (0.60–0.83)		0.61 (0.45–0.77)			
Prevalence of biopsy Gleason score (≥ 7)	4	≥ 75.4%[†]	0.74 (0.67–0.82)	0.278	0.71 (0.60–0.82)	0.78	88.20	< 0.01
	3	< 75.4%[†]	0.80 (0.73–0.88)		0.73 (0.62–0.85)			
Magnet strength	8	3T	0.78 (0.72–0.84)	0.589	0.68 (0.59–0.77)	0.811	0.54	0.76
	5	1.5T	0.81 (0.73–0.89)		0.66 (0.54–0.78)			
Experience of MR readers	6	≥ 10 years	0.79 (0.74–0.84)	0.613	0.64 (0.57–0.72)	0.016	52.61	< 0.01
	4	< 10 years	0.81 (0.75–0.88)		0.80 (0.72–0.88)			
Endorectal coil	5	Used	0.72 (0.66–0.77)	0.001	0.63 (0.51–0.75)	0.344	10.37	0.01
	8	Not used	0.83 (0.79–0.87)		0.70 (0.62–0.79)			
Publication year	11	After 2000	0.80 (0.75–0.85)	0.354	0.69 (0.62–0.76)	0.303	2.04	0.36
	2	Before 2000	0.72 (0.57–0.88)		0.58 (0.38–0.77)			

*10 mm as stated by PI-RADS v2 guideline in one study (26), and unclear explanation in two (31, 33), [†]Median value of all included studies. CI = confidence interval, LRT = likelihood-ratio test, MR = magnetic resonance

with neurovascular bundle sparing. On the other hand, high specificity could be favored when there is a need to avoid withholding potential curative treatment (11). Therefore, based on the results of our study, we believe that LCC on MRI can provide incremental value in the management of patients with prostate cancer, especially in clinical settings where high sensitivity for predicting EPE is needed.

Substantial heterogeneity existed among the included studies regarding MRI sequences, threshold values and how they were determined, and methods for measuring LCC. However, all sequences, including T2WI, ADC, and DCE, showed similar diagnostic performance (sensitivities and specificities of 0.76–0.80 and 0.67–0.68, respectively). Although not significantly different, the sensitivity and specificity (0.84 and 0.76, respectively) using the maximum value from all sequences was slightly greater in one study (16). This may be related to due fact that tumor measurements on MRI tend to underestimated compared with that performed using pathological specimens, and using the maximum value could potentially lessen the degree of MRI-based underestimation of LCC relative to the pathological LCC (36). However, there are also studies stating that MRI-based LCC correlates well with pathologic-LCC without any overestimation or underestimation (14, 37). Therefore, conclusions regarding the incremental value

of using the maximum value cannot be drawn as there are not enough data to support one and further studies are warranted.

Although different cutoff values ranging from 6 to 20 mm were used, this was not shown to affect the overall heterogeneity (p = 0.93 and 0.88 using cutoff values of 10 and 12 mm). Moreover, a threshold effect was not observed, which is generally expected in meta-analyses using thresholds for a continuous variable like LCC. Still, increasing LCC has been shown to be associated with a greater probability of EPE (30, 38). For example, in the study by Masumoto et al. (30, 38), every increase in 1 mm of LCC was associated with a 13% increase in the odds for EPE. In addition, there are studies that suggest anterior tumors are less aggressive and that different threshold values, specifically, a less strict criterion using higher cutoff values, should be used for anterior tumors compared with that used for posterior tumors (38, 39). Therefore, the cutoff value should be tailored to several factors, such as location of the tumor and the clinical setting (e.g., using a lower cutoff to detect EPE more sensitively). Curvilinear measurement of LCC on MRI, theoretically, may reflect pathological LCC better than linear measurement, and was a significant factor affecting heterogeneity. However, there was no significant difference in the sensitivity

or specificity; thus, it may seem reasonable to use any available ruler tool that is provided by the image viewing software.

Sensitivity and specificity from more experienced readers were not significantly better than those from less experienced readers were. At first, this may be unexpected, as it has been shown in the literature than the accuracy of EPE, using prostate MRI is dependent on experience (9, 40). However, those studies were based on subjective assessment on EPE while LCC can be considered more objective and quantitative, rendering it less dependent on experience and potentially enhancing the reproducibility between readers with different experience levels. In fact, substantial to almost perfect agreement (kappa values of 0.70–0.98 and intraclass correlation coefficients of 0.979–0.983) for measuring LCC was shown in the majority of studies assessing interreader agreement (14, 15, 29, 32, 33).

In this meta-analysis, technical aspects of MRI were investigated. Magnet strength (3- vs. 1.5T) did not have a statistically significant effect on either sensitivity (0.78 and 0.81 for 3- and 1.5T, respectively) or specificity (0.68 and 0.66, respectively). This contrasts with results of a previous meta-analysis assessing the performance of MRI for detecting EPE using subjective assessment (sensitivity of 0.61 and 0.55 for 3- and 1.5T, respectively) (8). Detection of subtle capsular irregularity or small foci of extracapsular tumor may require high spatial resolution, which is easier to obtain at higher magnetic strengths; however, it seems that LCC may be less affected by magnet strength when ≥ 1.5T scanners are used. Therefore, both 1.5- and 3T scanners may provide comparable and objective LCC measurement on MRI provided MRI protocols are optimized. Studies with endorectal coils showed significantly lower sensitivity compared with those without. We speculate that this may stem from the fact that endorectal coils can lead to deformation of the prostatic contour possibly influencing measurement of LCC (41).

In the present study, prevalence of high Gleason tumors was a factor affecting heterogeneity, but no significant differences were observed in either sensitivity or specificity estimates. Nevertheless, in one of the included studies that additionally evaluated LCC stratified by tumor grade, mean LCC was significantly larger for higher grade tumors (15.3 mm vs. 9.0 mm, $p = 0.0001$) (25). Furthermore, for the LCC criterion of < 10 mm, 41.6% of higher-grade tumors still had EPE compared with only 2.8% of lower grade tumors (25). Bakir et al. (37) recently stated pathology-based

LCC cutoff values decreased as the International Society of Urological Pathology grade group increased in terms of EPE positivity, substantiating the possible influence of tumor grade on the LCC-EPE relationship. Further studies are warranted to investigate whether different threshold values are required for tumors with different Gleason scores.

Our study has several limitations. Firstly, all included studies were performed retrospectively, which may have introduced selection bias. Second, as full organ histopathological correlation was required to evaluate pathological EPE, we solely included patients that underwent RP. Therefore, caution is needed when applying these results to a non-surgical population (i.e., radiotherapy, active surveillance, or recurrence). Third, the heterogeneity was substantial between the studies. Although we performed meta-regression analyses and sensitivity analyses to identify potential factors attributable to this heterogeneity, some factors remain unexplained. Fourth, although we were able to derive some conclusions, for instance, that LCC can be measured using any MRI sequence with either curvilinear or linear tools, the optimal threshold value of LCC could not be established. The cutoff value should be tailored to the likelihood of EPE, and the clinician and patient's preferences for management.

In conclusion, greater LCC on MRI was associated with a higher probability of EPE. Despite its overall moderate diagnostic performance, the relatively higher sensitivity compared with that of conventional subjective assessment might be of incremental value for helping select candidates for active surveillance or functional preservation treatments by avoiding underestimation of the disease. Furthermore, as LCC is relatively simple to measure and is less dependent on reader experience, it can be considered a reproducible and objective quantitative predictor in the assessment of EPE in prostate cancer. However, further studies are needed to establish the optimal cutoff value for each clinical setting.

Conflicts of Interest

For conflict of interest, since May 2017, Dr. Hricak has served on the Board of Directors of Ion Beam Applications (IBA), a publicly traded company, and she receives annual compensation for her service. Furthermore, Dr. Hricak is a member of the External Advisory Board of the University of Michigan Comprehensive Cancer Center, the International Advisory Board of the University of Vienna, Austria, and the Scientific Committee of the DKFZ (German Cancer Research Center), Germany; she does not receive financial

compensation for any of these roles. Otherwise, we do not have any other conflict of interest to disclose.

ORCID iDs
Sungmin Woo
https://orcid.org/0000-0001-8459-8369
Tae-Hyung Kim
https://orcid.org/0000-0002-0333-8475
Sangwon Han
https://orcid.org/0000-0001-8095-0396
Chong Hyun Suh
https://orcid.org/0000-0002-4737-0530
Soleen Ghafoor
https://orcid.org/0000-0002-8459-7363
Hedvig Hricak
https://orcid.org/0000-0003-2240-9694
Hebert Alberto Vargas
https://orcid.org/0000-0003-1169-133X

REFERENCES

1. Godoy G, Tareen BU, Lepor H. Site of positive surgical margins influences biochemical recurrence after radical prostatectomy. *BJU Int* 2009;104:1610-1614
2. Mikel Hubanks J, Boorjian SA, Frank I, Gettman MT, Houston Thompson R, Rangel LJ, et al. The presence of extracapsular extension is associated with an increased risk of death from prostate cancer after radical prostatectomy for patients with seminal vesicle invasion and negative lymph nodes. *Urol Oncol* 2014;32:26.e1-26.e7
3. Rosen MA, Goldstone L, Lapin S, Wheeler T, Scardino PT. Frequency and location of extracapsular extension and positive surgical margins in radical prostatectomy specimens. *J Urol* 1992;148(2 Pt 1):331-337
4. Boccon-Gibod L, Bertaccini A, Bono AV, Dev Sarmah B, Höltl W, Mottet N, et al. Management of locally advanced prostate cancer: a European consensus. *Int J Clin Pract* 2003;57:187-194
5. Boyce S, Fan Y, Watson RW, Murphy TB. Evaluation of prediction models for the staging of prostate cancer. *BMC Med Inform Decis Mak* 2013;13:126
6. Tsuzuki T, Hernandez DJ, Aydin H, Trock B, Walsh PC, Epstein JI. Prediction of extraprostatic extension in the neurovascular bundle based on prostate needle biopsy pathology, serum prostate specific antigen and digital rectal examination. *J Urol* 2005;173:450-453
7. Hoeks CM, Barentsz JO, Hambrock T, Yakar D, Somford DM, Heijmink SW, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology* 2011;261:46-66
8. de Rooij M, Hamoen EH, Witjes JA, Barentsz JO, Rovers MM. Accuracy of magnetic resonance imaging for local staging of prostate cancer: a diagnostic meta-analysis. *Eur Urol* 2016;70:233-245
9. Wibmer A, Vargas HA, Donahue TF, Zheng J, Moskowitz C, Eastham J, et al. Diagnosis of extracapsular extension of prostate cancer on prostate MRI: impact of second-opinion readings by subspecialized genitourinary oncologic radiologists. *AJR Am J Roentgenol* 2015;205:W73-W78
10. Hole KH, Axcrona K, Lie AK, Vlatkovic L, Geier OM, Brennhovd B, et al. Routine pelvic MRI using phased-array coil for detection of extraprostatic tumour extension: accuracy and clinical significance. *Eur Radiol* 2013;23:1158-1166
11. Woo S, Cho JY, Kim SY, Kim SH. Extracapsular extension in prostate cancer: added value of diffusion-weighted MRI in patients with equivocal findings on T2-weighted imaging. *AJR Am J Roentgenol* 2015;204:W168-W175
12. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic performance of prostate imaging reporting and data system version 2 for detection of prostate cancer: a systematic review and diagnostic meta-analysis. *Eur Urol* 2017;72:177-188
13. Baco E, Rud E, Vlatkovic L, Svindland A, Eggesbø HB, Hung AJ, et al. Predictive value of magnetic resonance imaging determined tumor contact length for extracapsular extension of prostate cancer. *J Urol* 2015;193:466-472
14. Onay A, Vural M, Armutlu A, Ozel Yıldız S, Kiremit MC, Esen T, et al. Evaluation of the most optimal multiparametric magnetic resonance imaging sequence for determining pathological length of capsular contact. *Eur J Radiol* 2019;112:192-199
15. 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;43:990-997
16. Woo S, Kim SY, Cho JY, Kim SH. Length of capsular contact on prostate MRI as a predictor of extracapsular extension: which is the most optimal sequence? *Acta Radiol* 2017;58:489-497
17. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM; the PRISMA-DTA Group. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;319:388-396
18. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-536
19. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982-990
20. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20:2865-2884
21. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin*

Epidemiol 2005;58:882-893

22. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-560

23. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions, Version 5.1.0 [updated March 2011]. Cochrane Collaboration, 2011. Available at: https://training.cochrane.org/handbook/archive/v5.1/. Accessed March 21, 2019

24. Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002;2:9

25. Caglic I, Povalej Brzan P, Warren AY, Bratt O, Shah N, Barrett T. Defining the incremental value of 3D T2-weighted imaging in the assessment of prostate cancer extracapsular extension. *Eur Radiol* 2019;29:5488-5497

26. Costa DN, Passoni NM, Leyendecker JR, de Leon AD, Lotan Y, Roehrborn CG, et al. Diagnostic utility of a Likert scale versus qualitative descriptors and length of capsular contact for determining extraprostatic tumor extension at multiparametric prostate MRI. *AJR Am J Roentgenol* 2018;210:1066-1072

27. Granja MF, Pedraza CM, Flórez DC, Romero JA, Palau MA, Aguirre DA. Predicting extracapsular involvement in prostate cancer through the tumor contact length and the apparent diffusion coefficient. *Radiologia* 2017;59:313-320

28. Kongnyuy M, Sidana A, George AK, Muthigi A, Iyer A, Ho R, et al. Tumor contact with prostate capsule on magnetic resonance imaging: a potential biomarker for staging and prognosis. *Urol Oncol* 2017;35:30.e1-30.e8

29. Krishna S, Lim CS, McInnes MDF, Flood TA, Shabana WM, Lim RS, et al. Evaluation of MRI for diagnosis of extraprostatic extension in prostate cancer. *J Magn Reson Imaging* 2018;47:176-185

30. Matsuoka Y, Ishioka J, Tanaka H, Kimura T, Yoshida S, Saito K, et al. Impact of the prostate imaging reporting and data system, version 2, on MRI diagnosis for extracapsular extension of prostate cancer. *AJR Am J Roentgenol* 2017;209:W76-W84

31. Mehralivand S, Shih JH, Harmon S, Smith C, Bloom J, Czarniecki M, et al. A grading system for the assessment of risk of extraprostatic extension of prostate cancer at multiparametric MRI. *Radiology* 2019;290:709-719

32. Outwater EK, Petersen RO, Siegelman ES, Gomella LG, Chernesky CE, Mitchell DG. Prostate carcinoma: assessment of diagnostic criteria for capsular penetration on endorectal coil MR images. *Radiology* 1994;193:333-339

33. Yu KK, Hricak H, Alagappan R, Chernoff DM, Bacchetti P, Zaloudek CJ. Detection of extracapsular extension of prostate carcinoma with endorectal and phased-array coil MR imaging: multivariate feature analysis. *Radiology* 1997;202:697-702

34. Van Holsbeeck A, Degroote A, De Wever L, Vanhoutte E, De Keyzer F, Van Poppel H, et al. Staging of prostatic carcinoma at 1.5-T MRI: correlation of a simplified MRI exam with whole-mount radical prostatectomy specimens. *Br J Radiol* 2016;89:20160101

35. Cornud F, Rouanne M, Beuvon F, Eiss D, Flam T, Liberatore M, et al. Endorectal 3D T2-weighted 1mm-slice thickness MRI for prostate cancer staging at 1.5Tesla: should we reconsider the indirects signs of extracapsular extension according to the D'Amico tumor risk criteria? *Eur J Radiol* 2012;81:e591-e597

36. Bratan F, Melodelima C, Souchon R, Hoang Dinh A, Mège-Lechevallier F, Crouzet S, et al. How accurate is multiparametric MR imaging in evaluation of prostate cancer volume? *Radiology* 2015;275:144-154

37. Bakir B, Onay A, Vural M, Armutlu A, Yıldız SÖ, Esen T. Can extraprostatic extension be predicted by tumor-capsule contact length in prostate cancer? Relationship with International Society of Urological Pathology grade groups. *AJR Am J Roentgenol* 2020;214:588-596

38. Matsumoto K, Akita H, Narita K, Hashiguchi A, Takamatsu K, Takeda T, et al. Prediction of extraprostatic extension by MRI tumor contact length: difference between anterior and posterior prostate cancer. *Prostate Cancer Prostatic Dis* 2019;22:539-545

39. Rud E, Diep L, Baco E. A prospective study evaluating indirect MRI-signs for the prediction of extraprostatic disease in patients with prostate cancer: tumor volume, tumor contact length and tumor apparent diffusion coefficient. *World J Urol* 2018;36:629-637

40. Mullerad M, Hricak H, Wang L, Chen HN, Kattan MW, Scardino PT. Prostate cancer: detection of extracapsular extension by genitourinary and general body radiologists at MR imaging. *Radiology* 2004;232:140-146

41. Heijmink SW, Scheenen TW, van Lin EN, Visser AG, Kiemeney LA, Witjes JA, et al. Changes in prostate shape and volume and their implications for radiotherapy after introduction of endorectal balloon as determined by MRI at 3T. *Int J Radiat Oncol Biol Phys* 2009;73:1446-1453