BJR

https://doi.org/10.1259/bjr.20191050

Received: 17 December 2019 | Revised: 08 July 2020 | Accepted: 09 September 2020

Cite this article as:
Wadera A, Alabousi M, Pozdnyakov A, Kashif Al-Ghita M, Jafri A, McInnes MDF, et al. Impact of PI-RADS Category 3 lesions on the diagnostic accuracy of MRI for detecting prostate cancer and the prevalence of prostate cancer within each PI-RADS category: A systematic review and meta-analysis. *Br J Radiol* 2020; **94**: 20191050.

SYSTEMATIC REVIEW

Impact of PI-RADS Category 3 lesions on the diagnostic accuracy of MRI for detecting prostate cancer and the prevalence of prostate cancer within each PI-RADS category: A systematic review and meta-analysis

[1]AKSHAY WADERA, [1]MOSTAFA ALABOUSI, [2]ALEX POZDNYAKOV, [3]MOHAMMED KASHIF AL-GHITA, [4]ALI JAFRI, [5,6]MATTHEW DF MCINNES, [6]NICOLA SCHIEDA, [7]CHRISTIAN B VAN DER POL, [8]JEAN-PAUL SALAMEH, [1]LUCY SAMOILOV, KAELA GUSENBAUER and [9]ABDULLAH ALABOUSI, MD, FRCPC

[1]Department of Radiology, McMaster University, Hamilton, ON, Canada
[2]Faculty of Medicine, McMaster University, Hamilton, ON, Canada
[3]Faculty of Biomedical Sciences, Western University, London, ON, Canada
[4]Department of Medicine, New York Institute of Technology School of Osteopathic Medicine, Glen Head, NY, United States
[5]The Ottawa Hospital Research Institute, Clinical Epidemiology Program, Ottawa, ON, Canada
[6]Department of Radiology, University of Ottawa, The Ottawa Hospital, Ottawa, ON, Canada
[7]Department of Radiology, McMaster University, Juravinski Hospital, Hamilton, ON, Canada
[8]Department of Medicine, Clinical Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[9]Department of Radiology, McMaster University, St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada

Address correspondence to: Dr Abdullah Alabousi
E-mail: *abdullah.alabousi@medportal.ca*

Objective: To evaluate Prostate Imaging Reporting and Data System (PI-RADS) category 3 lesions' impact on the diagnostic test accuracy (DTA) of MRI for prostate cancer (PC) and to derive the prevalence of PC within each PI-RADS category.

Methods: MEDLINE and Embase were searched until April 10, 2020 for studies reporting on the DTA of MRI by PI-RADS category. Accuracy metrics were calculated using a bivariate random-effects meta-analysis with PI-RADS three lesions treated as a positive test, negative test, and excluded from the analysis. Differences in DTA were assessed utilizing meta-regression. PC prevalence within each PI-RADS category was estimated with a proportional meta-analysis.

Results: In total, 26 studies reporting on 12,913 patients (4,853 with PC) were included. Sensitivities for PC in the positive, negative, and excluded test groups were 96% (95% confidence interval [CI] 92–98), 82% (CI 75-87), and 95% (CI 91-97), respectively. Specificities for the positive, negative, and excluded test groups were 33% (CI 23-44), 71% (CI 62-79), and 52% (CI 37-66), respectively. Meta-regression demonstrated higher sensitivity ($p < 0.001$) and lower specificity ($p < 0.001$) in the positive test group compared to the negative group. Clinically significant PC prevalences were 5.9% (CI 0-17.1), 11.4% (CI 6.5–17.3), 24.9% (CI 18.4–32.0), 55.7% (CI 47.8–63.5), and 81.4% (CI 75.9–86.4) for PI-RADS categories 1, 2, 3, 4 and 5, respectively.

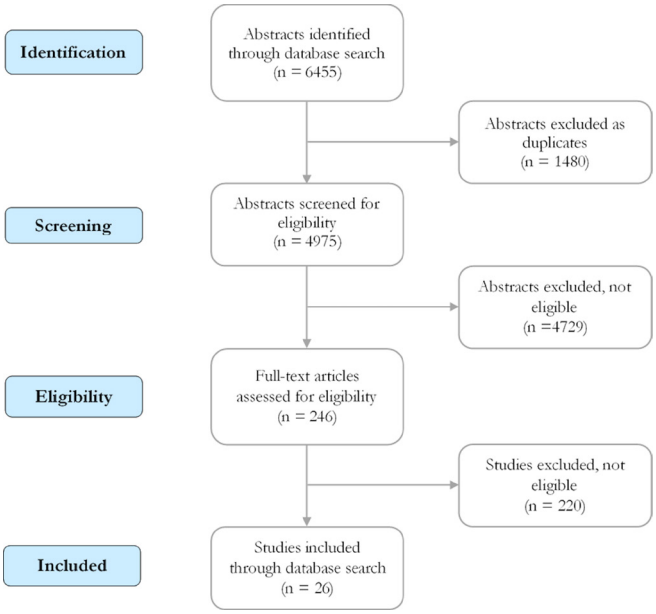
Conclusion: PI-RADS category 3 lesions can significantly impact the DTA of MRI for PC detection. A low prevalence of clinically significant PC is noted in PI-RADS category 1 and 2 cases.

Advances in knowledge: Inclusion or exclusion of PI-RADS category 3 lesions impacts the DTA of MRI for PC detection.

INTRODUCTION

Prostate cancer (PC) is one of the leading causes of death among males in the United States and Western Europe.[1] Prostate cancer alone accounts for almost one in five new cancer diagnoses, and the risk of developing invasive PC is approximately one in nine.[2] The gold-standard for PC diagnosis involves the use of transrectal ultrasound (TRUS)-guided biopsies with systematic sampling of the prostate gland in the context of increased clinical suspicion, including an elevated serum prostate-specific antigen (PSA) and/or an abnormal digital rectal exam (DRE).[3,4] Histopathology from prostate biopsies are often reported using the Gleason Score, a grading system for tumor aggressiveness.[5,6] This practice has been shown to reduce mortality.[7]

Figure 1. Study flow diagram.



MRI has emerged as an important diagnostic test to assess for clinically significant PC.[8] The most frequently utilized protocol, multiparametric MRI (mpMRI), includes T_2 weighted imaging (T_2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE) sequences.[8] Interpretation and reporting of mpMRI is based on the Prostate Imaging-Reporting and Data System (PI-RADS), originally introduced in 2012 and most recently revised in v. 2.1 in 2019, which utilizes a 5-point Likert scale indicating the probability that a lesion represents a clinically significant PC.[9,10] PI-RADS categories of very low and low (1 and 2) and high and very high (4 and 5) likelihood of clinically significant PC are commonly treated as "negative" and "positive" test results in diagnostic test accuracy studies, respectively.[10] A PI-RADS category 3 result, for which the risk of clinically significant cancer is "equivocal", presents a diagnostic challenge, as it is treated as a "positive" result in some studies[11] and "negative" in others.[12] Furthermore, the number of patients classified with a PI-RADS category 3 lesion on mpMRI is considerable, varying between one in three and one in five,[13] indicating a need to further explore their impact on the diagnostic accuracy of mpMRI.

In this context, the diagnostic accuracy of mpMRI may be influenced by the threshold selected for a "positive" test result, which may in turn limit the between-study comparisons for mpMRI. Thus, our objective was to investigate the impact of different PI-RADS category thresholds on the diagnostic test accuracy of mpMRI for the detection of PC. Our hypothesis was that the variable classification of PI-RADS category 3 lesions is associated with significant differences in the diagnostic test accuracy of mpMRI. If true, a standardized threshold may be warranted to reduce this variability. A secondary objective was to assess the prevalence of reported PC within each PI-RADS category.

METHODS AND MATERIALS

Literature search

A protocol for this study was registered on the Open Science Framework (osf.io/czb9n). We performed a literature search of electronic databases Medline and Embase to identify all relevant studies published until April 10, 2020. We limited the search to studies published on January 1, 2012 or later, as this was the year of publication of the first PI-RADS guidelines.[14] Details of the search strategy, created in consultation with a librarian, are included in Supplementary Material 1.The references sections of the included studies were manually searched to identify additional studies for inclusion.

Eligibility criteria and study selection

Inclusion criteria were defined as follows: (i) the diagnostic tests reported on accuracy of mpMRI in patients with suspected PC; (ii) studies that included treatment-naïve patients with or without prior prostate biopsy; (iii) the reference standard was histopathology from prostate biopsies or prostatectomy specimens; (iv) the results report sufficient per-patient or per-lesion data to construct a 2 × 2 contingency table and to stratify each true positive, false negative, true negative, and false positive by each individual PI-RADS category; (v) the full text was available in English.

Exclusion criteria were defined as follows: (i) the study investigated diagnostic test accuracy only in post-treatment patients (surgery and/or focal therapy); (ii) the study was a review, commentary, case report, case series or letter to the editor. For duplicate publications, the study with the largest sample size was included.

Results of the literature search were imported into a reference management software (Reference Manager 11, 2008; Thomson Reuters, Toronto, ON, Canada) for independent title and abstract review (Phase I) followed by independent full-text screening. Discrepancies were resolved by consensus.

Data extraction

Data extraction was performed independently using the included studies by multiple investigators. Investigators performed double blinded data extraction of the first five studies to improve familiarity and consistency. Discrepancies were resolved by consensus. The following data metrics were extracted into a spreadsheet program (Microsoft Excel, 2016; Microsoft, Redmond, WA) using predefined forms: first author, study title, year and journal of publication, country of corresponding author, study design, patient demographics, PI-RADS version, technical imaging characteristics, reference standard specifications, and 2 × 2 contingency table data (true positives, false negatives, true negatives, and false positives) stratified according to PI-RADS category.

Data analysis

We utilized three groups for our analysis: the "positive test" group where PI-RADS category 3 was treated as a positive mpMRI result for clinically significant PC; the "negative test" group where PI-RADS category 3 was treated as a negative mpMRI result; and the "excluded test" group where PI-RADS

Table 1. Characteristics of included studies

Study	Design	Number of patients/ lesions	No. of PC diagnoses	Age, years (mean or median)	Gleason score threshold	PI-RADS version	PSA (mean or median)	MRI Tesla strength	Number of interpreters	Endorectal Coil	Biopsy or prostatectomy
Ahmed et al. 2017	Prospective	576	230	63	≥7	2	7	1.5	2	No	Biopsy
Boesen et al. 2018	Prospective	289	88	64	≥7	1	12	3	1	No	Biopsy
Jordan et al. 2017	Retrospective	282	81	67	≥7	2	9	3	1	No	Biopsy
Toner et al. 2017	Retrospective	152	92	63	≥7	2	6	1.5 or 3	1	NR	Prostatectomy
Thompson et al. 2016	Retrospective	344	141	63	≥7	2	5	1.5 or 3	2	NR	Biopsy
Lista et al. 2015	Prospective	150	28	66	≥6	2	11	1.5	NR	Yes	Biopsy
Zhao et al. 2016	Retrospective	372	155	69	≥7	2	15	3	2	Yes	Biopsy
Grey et al. 2015	Prospective	201	77	65	≥6	2	13	1.5	1	No	Biopsy
Numao et a. 2013	Prospective	351	126	65	≥7	NR	6	1.5	1	No	Biopsy
Osses et al. 2016	Retrospective	156	63	68	≥7	1	11	3	2	No	Mixed
Feng et al. 2016	Retrospective	401	150	64	NR	Both	44	3	2	Yes	Biopsy
Salami et al. 2014	Prospective	175	83	65	≥7	2	7	3	3	Yes	Mixed
Thompson et al. 2014	Prospective	150	51	62	≥7	2	6	1.5 or 3	2	No	Mixed
Greer et al. 2017	Retrospective	268	244	62	≥7	2	7.09	3	9	Yes	Prostatectomy
Mussi et al. 2017	Retrospective	118	48	NR	≥6	2	4.6	3	2	No	Biopsy
Choi et al. 2019	Retrospective	113	84	65	≥7	2	7.9	3	2	NR	Prostatectomy
Duan et al. 2019	Retrospective	231	58	65	≥7	2	6.8	NR	1	NR	Biopsy
Donato et al. 2020	Retrospective	344	208	65	≥7	Both	6.8	3	5	No	Biopsy
Gaur et al. 2018	Prospective	733	236	64	≥7	2	6.5	3	1	Both	Biopsy
Hsieh et al. 2020	Prospective	102	24	65	≥7	2	7.78	3	1	No	Biopsy
Lee et al. 2018	Retrospective	237	106	65	≥7	2	9.7	1.5 or 3	1	No	biopsy
Luzzago et al. 2018	Prospective	250	117	63	≥7	Both	6.1	1.5	3	No	Mixed
Pal et al. 2018	Retrospective	426	196	64	≥7	1	6.2	1.5	2	No	Biopsy
Rozas et al. 2019	Retrospective	342	83	NR	≥7	2	NR	3	2	No	Biopsy
Viana et al. 2018	Retrospective	98	38	60	≥6	2	6.3	3	2	No	Biopsy
Zhang et al. 2018(1)	Prospective	114	39	66	≥7	1	9.7	3	2	No	Biopsy

(Continued)

Table 1. (Continued)

Study	Design	Number of patients/ lesions	No. of PC diagnoses	Age, years (mean or median)	Gleason score threshold	PI-RADS version	PSA (mean or median)	MRI Tesla strength	Number of interpreters	Endorectal Coil	Biopsy or prostatectomy
Zhang et al. 2018	Prospective	123	67	66	≥7	1	11.1	3	2	No	Biopsy
Westphalen et al. 2020	Retrospective	3,449	2,082	65	≥7	2	6.6	1.5 or 3	NR	No	Biopsy

PC, prostate cancer;PI-RADS, Prostate Imaging Reporting and Data System; PSA, prostate specificantigen; NR, not reported.

category 3 lesions were entirely excluded from the diagnostic accuracy analysis. A bivariate random effects model was used to pool data and generate summary estimates for sensitivity and specificity according to each of the positive test, negative test, and excluded test groups.[15,16] Forest plots and hierarchical summary receiver operator characteristic (hsROC) curves were constructed. Area under the curve (AUC) mean estimates were calculated. Comparison of sensitivity and specificity for each of the test groups was performed with comparative meta-regression models. A p-value < 0.05 was considered statistically different within the meta-regression model. Additional subgroup analysis to identify potential sources of heterogeneity and risk of bias assessment were not performed, as this has been previously explored.[17] Proportional meta-analyses were performed to determine estimates of the prevalence of any PC (Gleason score ≥6), as well as a sensitivity analysis to estimate the prevalence of clinically significant PC (Gleason score ≥7), for each individual PI-RADS categories using a random effects model with arcsine transformation.[16,18,19] The prevalence of PC for combined PI-RADS categories 1 & 2 was assessed as well, as some studies only reported the combined results of these categories. Forest plots were created using the estimated model parameters. Heterogeneity was assessed using the I^2 value, with valuesgreater than 50% considered at risk for substantial variability. Analysis was performed using the "metaprop"and "meta" packages in STATA v. 11.2 (Texas, United States) and R v. 3.5.1 (Vienna, Austria).[18,19]

RESULTS

A study flow diagram is shown in Figure 1. An initial 4,975 studies underwent title and abstract screening, of which 246 studies were retrieved for full text review. Reasons for exclusion of studies included the following: stratified 2 × 2 contingency table data stratified according to PI-RADS category not provided; the included study did not report per-patient analysis (*i.e.* only per-lesion analysis); the included patients had previous treatment or intervention. In all, 26 studies reporting on 12,913 patients/lesions (4,853 with PC) met the inclusion criteria for meta-analysis.[11,12,20–43] Table 1 provides a summary of the included studies. Of the 26 studies, 18 studies reported PI-RADS 3 as a positive test result in their analysis.[12,20–23,25,27–30,32–38,42] The remaining eight studies reported PI-RADS 3 lesions as negative.[11,24,26,31,39–41,43] The median age range for all included studies was 62–68 years.

A Gleason score threshold of ≥7 was used for 23 studies.[11,12,20–26,29–41,43] Four studies reported data on lesions with a Gleason score threshold of ≥6.[11,27,28,42] One study did not explicitly report the Gleason score threshold for PC.[11] PI-RADS version 2 was used for 22 studies.[11,12,20,21,23–25,27–35,37–39,41–43] Four studies strictly used PI-RADS version 1,[22,26,36,40] while three studies used both versions 1 and 2.[11,20,35] No study used the newly released PI-RADS version 2.1.[10] PC was confirmed by pathology using either biopsy or prostatectomy, with two studies strictly using prostatectomy results.[41,43]

Table 2 summarizes the diagnostic test accuracy results for the positive test, negative test, and excluded test groups. Figure 2 shows the pooled sensitivity and specificity for the positive,

Table 2. Diagnostic test accuracy according to PI-RADS 3 category grouping

	PT Group (PI-RADS 3 = positive test)	NT Group (PI-RADS 3 = negative test)	ET Group (PI-RADS 3 = Excluded)
Sensitivity	96% (CI 92–98)	82% (CI 75–87)	95% (CI 91–97)
Specificity	33% (CI 23–44)	71% (CI 62–79)	52% (37-66)
AUC	0.81	0.84	0.89

AUC, area under the curve; PI-RADS, Prostate Imaging Reporting and Data System.
PT Group: PI-RADS 3, 4, 5 = positive test; PI-RADS 1 & 2 = negative test; NT Group: PI-RADS 4 & 5 = positive test; PI-RADS 1, 2, 3 = negative test; ET Group: PI-RADS 4 & 5 = positive test; PI-RADS 1 & 2 = negative test.

negative, and excluded test groups. Sensitivities for the positive, negative, and excluded test groups were 96% (95% confidence interval [CI] 92–98), 82% (CI 75–87), and 95%

(CI 91–97), respectively. Specificities for the positive, negative, and excluded test groups were 33% (CI 23–44), 71% (CI 62–79), and 52% (CI 37–66), respectively. Figure 3 illustrates the comparative summary ROC curves for the positive test (AUC = 0.81), negative test (AUC = 0.84), and excluded test (AUC = 0.89) groups.

Table 3 provides a summary of the comparative multivariate meta-regression models for the positive, negative, and excluded test groups. Within the first model, the excluded test group was used as the reference for comparison to the positive and negative test groups. The positive test group demonstrated a statistically significant lower specificity ($p = 0.022$), while the sensitivity was no different than the excluded test group ($p = 0.598$). Meanwhile, the negative test group a statistically significant higher specificity ($p = 0.030$) and lower sensitivity ($p < 0.001$) compared to the excluded group. The second model used the negative test group as a reference for comparison to the positive test group. The positive test group demonstrated a statistically significant lower specificity ($p < 0.001$), and higher sensitivity ($p < 0.001$) than the negative test group. The findings of the meta-regression model were compatible with the unadjusted pooled estimates of the mean for diagnostic accuracy of the positive, negative, and excluded test groups.

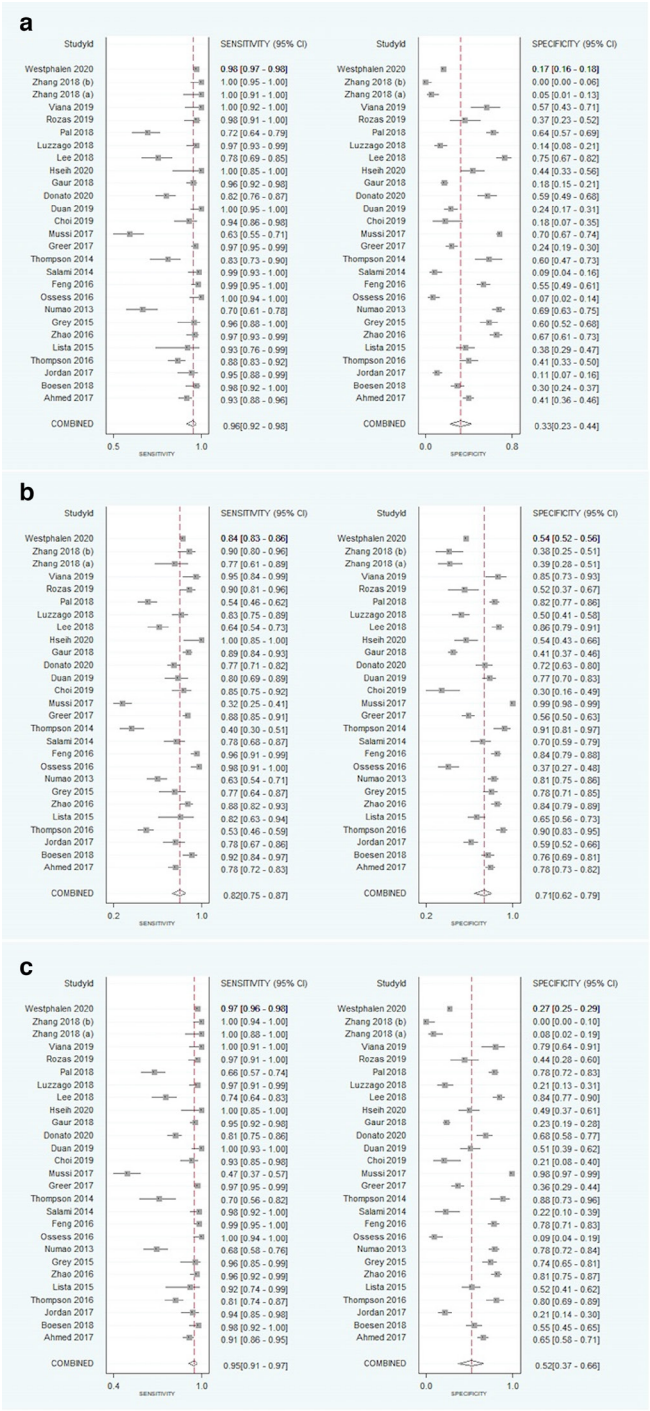
Figure 4 demonstrates forest plots and estimates of the prevalence of PC within each of the following categories: (a) PI-RADS category 1; (b) PI-RADS category 2; (c) combined PI-RADS categories 1 and 2; (d) PI-RADS category 3; (e) PI-RADS category 4; and (f) PI-RADS category 5. Figure 5 illustrates a flow diagram of patients/lesions included according to PI-RADS category classification. PI-RADS category 1 included a total of 396 patients (52 with PC), for which the pooled estimate of cancer prevalence was 3.4% (CI 0–11.6, I^2 =1 74.6%). PI-RADS category 2 included a total of 2,365 patients/lesions (270 with PC) for which the pooled estimate of cancer prevalence was 10.5% (CI 6.0–15.1, $I^{2=}$86.0%). Combined PI-RADS categories 1–2 included a total of 2,975 patients/lesions (350 with PC) for which the pooled estimate of cancer prevalence was 9.7% (CI 6.2–13.9, I^2 = 88.7%). PI-RADS category 3 included a total of 3,282 patients/lesions (662 with PC) for which the pooled estimate of cancer prevalence was 23.5%(CI 18.0–29.6, I^2 = 90.8%). PI-RADS category 4 included a total of 4,217 patients/lesions (1,952 with PC), for which the pooled estimate of cancer prevalence was 55.7% (CI 48.3–68.0, I^2 = 94.1%). PI-RADS category 5 included a total of 2,439 patients/lesions (1,889 with PC), for which the pooled estimate of cancer prevalence was 81.0% (CI 75.8–85.8, I^2 = 86.8%). Sensitivity analysis to specifically assess the prevalence of clinically significant PC did not result in decreased overall prevalence of PC in each PI-RADS category: PI-RADS 1 was 5.9% (CI 0–17.1,I^2 = 71.0%); PI-RADS 2 was 11.4% (CI 6.5–17.3, I^2 = 87.7%); combined PI-RADS 1–2 was 11.4% (CI6.9–16.8, I^2 = 89.4%); PI-RADS 3 was 24.9% (CI18.4–32.0, I^2 = 92.1%); PI-RADS 4 was 55.7% (CI47.8–63.5, I^2 = 94.7); and PI-RADS 5 was 81.4% (CI75.9–86.4, I^2 = 87.4%). Substantial variability was present for each of the PI-RADS categories.

DISCUSSION

This systematic review compared the diagnostic test accuracy estimates for mpMRI using different PI-RADS threshold values for the detection of PC in treatment-naïve patients utilizing 26 studies, reporting on 12,913 patients/lesions (4,853 with PC). Our study findings indicated that PI-RADS category 3 lesions can significantly impact the diagnostic accuracy measures of mpMRI, and that PI-RADS category 3 lesions are treated variably across different institutions and studies. Eighteen included studies designated category 3 lesions as a "positive" result, while the remaining 8 studies designated category 3 lesions as a "negative" result. And, although there is no set standard on the categorization of PI-RADS 3 lesion, the results of these heterogeneous studies can significantly influence clinical practice pertaining to prostate MRI. Furthermore, the number of PI-RADS category 3 lesions was not negligible, making up to one-quarter of the total number of all lesions identified. Of these lesions, clinically significant PC was identified in almost one-quarter of them. In the context of these findings, we believe that the reporting of individual PI-RADS categories in diagnostic accuracy studies is warranted, with an associated positive predictive value for each category. Furthermore, a standardized grouping method may be considered for the calculation of summary accuracy measures of mpMRI using PI-RADS.

One potential grouping method to address this impact on diagnostic accuracy is to completely exclude PI-RADS 3 lesions from data analyses. PI-RADS 3 lesions are considered "equivocal" for clinically significant PC, meaning that clinically they are neither treated as positive or negative. These patients will likely be followed closely with imaging or undergo additional biopsies. Therefore, the risk of missing a

Figure 2. Forest plots of the pooled sensitivity (i) and specificity (ii) in using PI-RADS Categoryto predict PC. Forest plots demonstrating PI-RADS three threshold as (A) positive, (B) negative, and (C) excluded. PI-RADS, Prostate Imaging Reporting and Data System.



clinically significant cancer diagnosis is likely low. However, this may not be practical as excluding these patients from data analysis of diagnostic test accuracy studies means excluding almost one-quarter of patients and a significant portion of cancer diagnoses.[13]

Figure 3. SROC curves for the diagnostic performance of PI-RADS using PI-RADS 3 threshold as (A) positive, (B) negative, and (C) excluded. PT: positive test; NT: negative test; ET: excluded test. PI-RADS, Prostate Imaging Reporting and Data System; SROC, summary receiver operating characteristic

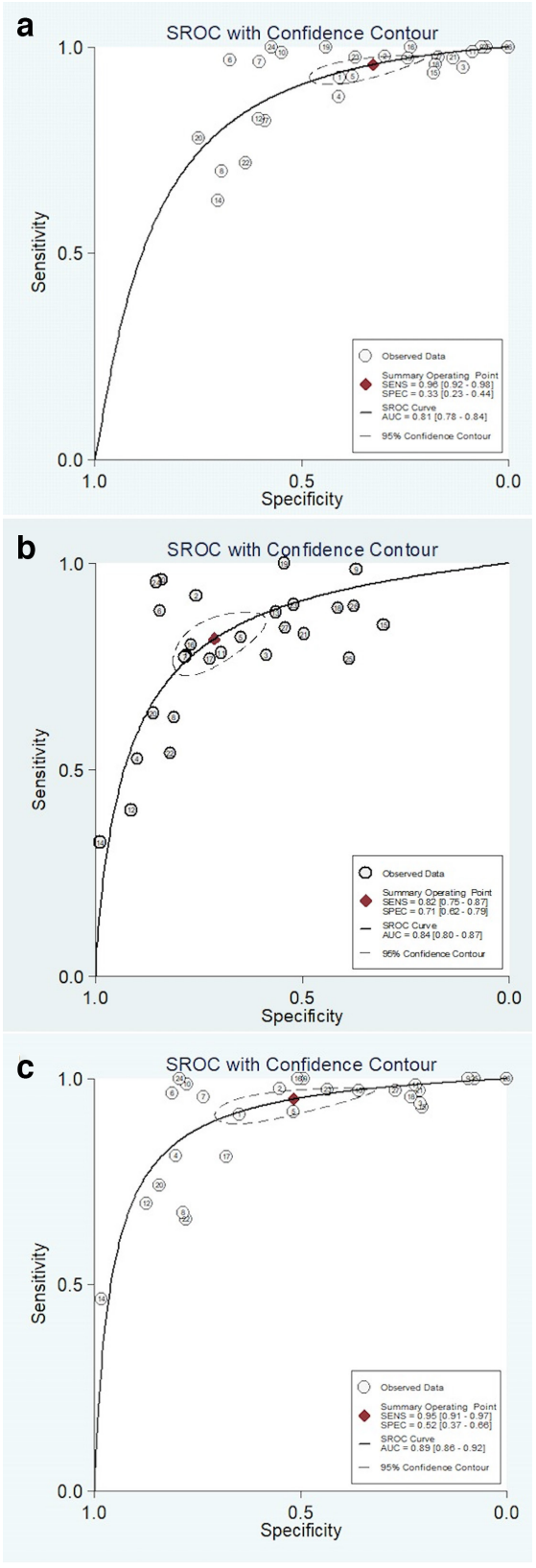


Table 3. Comparative multivariate meta-regression model comparing the diagnostic accuracy of the positive, negative, and excluded groups

Covariate	β Coefficient (95% CI)	Standard error	*P*-value
Sensitivity			
Test Group: Positive – reference excluded	0.174 (-0.473;0.821)	0.330	0.598
Test Group: Negative – reference excluded	−1.136 (-1.763;−0.510)	0.320	<0.001[a]
Test Group: Positive – reference negative	1.302 (0.715;1.889)	0.300	<0.001[a]
Specificity			
Test Group: Positive – reference excluded	0.794 (0.113;1.475)	0.347	0.022[a]
Test Group: Negative – reference excluded	−0.750 (-1.428;−0.073)	0.346	0.030[a]
Test Group: Positive – reference negative	1.531 (0.922;2.139)	0.311	<0.001[a]

PI-RADS, Prostate Imaging Reporting and Data System.
[a]denotes statistically significant result (p< 0.05)

Another potential grouping method is to select a standardized threshold. Our analysis indicates that treating PI-RADS 3 lesions as positive lowers specificity, while treating these lesions as negative will increase specificity but lower sensitivity. Moreover, selecting a PI-RADS threshold of category 3 or greater as a "positive" test result would disregard the utility of the DCE sequence altogether. The benefit of the DCE sequence is that it may "upgrade" peripheral zone PI-RADS 3 lesions to P-RADS 4 if it is positive.[10] However, if all PI-RADS 3 lesions were treated as a "positive" test results, the DCE findings would not affect the diagnostic accuracy. As a result, treating all PI-RADS 3 lesions as positive would limit our ability to compare the accuracy of mpMRI *vs* biparametric MRI (bpMRI) protocols, which only utilize the T_2WI and DWI sequences.[17] Based on these limitations, we believe all PI-RADS category 3 lesions should be treated as a "negative" test result for mpMRI diagnostic accuracy studies, as this will allow for the assessment of the utility of DCE, as well as allow for comparison to bpMRI.[17]

Another option could be to correlate with the clinical context. For instance, in biopsy-naïve patients, it may be beneficial to consider PI-RADS 3 studies as negative with the option for imaging follow-up or baseline biopsy. On the other hand, in patients with a prior positive biopsy or prior negative biopsy and persistent high clinical suspicion of cancer, a PI-RADS 3 lesion can be considered positive. However, this would require further study.

Figure 4. Forest plots and estimates of PC prevalence for the following categories: (A) PI-RADS 1; (B) PI-RADS 2; (C) combined PI-RADS 1 and 2; (D) PI-RADS 3; (E) PI-RADS 4; and (F) PI-RADS 5. PC, prostate cancer; Pi-RADS, Prostate Imaging Reporting and Data System.

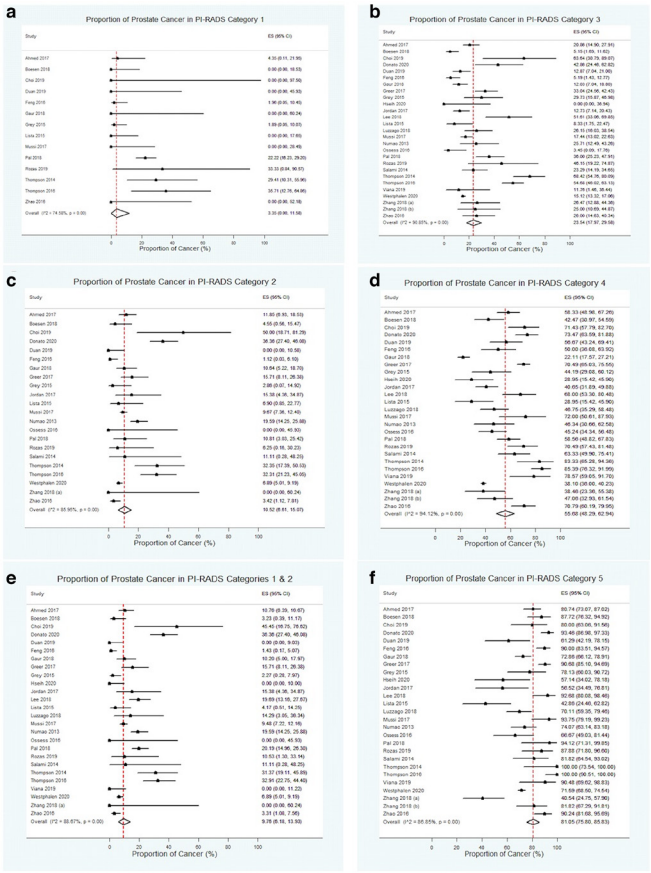
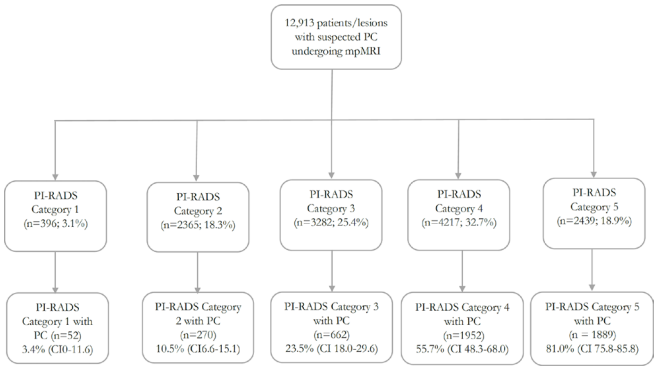


Figure 5. Flow diagram with stratification of patients/lesions by PI-RADS category. mpMRI: multiparametric MRI; PC: prostate cancer; PI-RADS: Prostate Imaging Reporting and Data System.



Previous studies have shown that the actual prevalence of clinically significant PC after targeted biopsy in PI-RADS 3 lesions vary between patients groups from 16 to 21%.[13] In comparison, our analysis indicated an even higher rate of any and clinically significant PC in PI-RADS category 3 lesions (24–25%). Furthermore, our findings demonstrated up to 5.9% of PI-RADS category 1 and 11.4% of PI-RADS category 2 lesions may demonstrate clinically significant PC. Considering category 1 and 2 lesions are not recommended to undergo targeted biopsy, this could contribute to one missed cancer for every 17 prostate MRIs classified as PI-RADS category 1 and one missed cancer for every nine prostate MRIs classified as PI-RADS category 2. Meanwhile, our pooled results indicate that over half of PI-RADS category 4 lesions and four of every five PI-RADS category 5 lesions are expected to be positive for PC, which is more congruent with expectations.

The findings of this study should be interpreted with caution as there are several limitations. First, our diagnostic accuracy analysis using the excluded test group is limited as it excludes a large proportion of PI-RADS lesions and cancer diagnoses. Secondly, there were several sources of heterogeneity between studies which may have masked some underlying differences; these have been previously assessed, including study design and technical MRI characteristics.[17] Furthermore, our search strategy did not include an assessment of the grey literature and non-English studies.

CONCLUSION

In summary, our analysis found that treating PI-RADS 3 lesions as a positive *vs* negative test result can significantly impact the diagnostic test accuracy of mpMRI in the detection of PC, which can ultimately influence clinical practice. Based on these findings, we believe standard reporting of individual PI-RADS categories and associated positive predictive values may be warranted. In our analysis, up to one-quarter of PI-RADS category 3 lesions represented PC. Moreover, clinically significant PC was found in up to 5.9% of PI-RADS category 1, 11.4% of PI-RADS category 2, and 24.9% of PI-RADS category 3 lesions, highlighting the importance of acknowledging that very low, low and equivocal likelihood of PC in PI-RADS categories 1, 2 and 3 lesions represent a non-zero risk of PC.

ACKNOWLEDGEMENTS

We would like to thank Stephanie Sanger, a McMaster University librarian, for her assistance with creating the search strategy for this study.

COMPETING INTERESTS

None to declare

FUNDING

This study was note funded.

ETHICS APPROVAL

Ethics approval was not required for this study as all data was available in the public domain.

DISCLOSURE

The authors of this study have no conflicts of interest to disclose.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;. ; **68**: 394–4242018. doi: https://doi.org/10.3322/caac.21492
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;. ; **66**: 7–302007. doi: https://doi.org/10.3322/caac.21332
3. Heijnsdijk EAM, Bangma CH, Borràs JM, de Carvalho TM, Castells X, Eklund M, et al. Summary statement on screening for prostate cancer in Europe. *Int J Cancer* 2018; **142**: 741–6. doi: https://doi.org/10.1002/ijc.31102
4. Mulhem E, Beaumont W, Heights S, Fulbright MN, Hospital P, Lyon S, et al. Prostate cancer screening. 2015;.
5. Gleason DF. Histologic grading of prostate cancer: a perspective. *Hum Pathol* 1992; **23**: 273–9. doi: https://doi.org/10.1016/0046-8177(92)90108-F
6. Gejerman G, Ciccone P, Goldstein M, Lanteri V, Schlecker B, Sanzone J, et al. Us preventive services Task force prostate-specific antigen screening guidelines result in higher Gleason score diagnoses. *Investig Clin Urol* 2017; **58**: 423. doi: https://doi.org/10.4111/icu.2017.58.6.423
7. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009; **360**: 1320–8. doi: https://doi.org/10.1056/NEJMoa0810084
8. Barentsz JO, Weinreb JC, Verma S, Thoeny HC, Tempany CM, Shtern F, et al. Synopsis of the PI-RADS V2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. *Eur Urol* 2016; **69**: 41–9. doi: https://doi.org/10.1016/j.eururo.2015.08.038
9. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* 2016; **69**: 16–40. doi: https://doi.org/10.1016/j.eururo.2015.08.052
10. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, 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**: 340–51. doi: https://doi.org/10.1016/j.eururo.2019.02.033
11. Feng Z-Y, Wang L, Min X-D, Wang S-G, Wang G-P, Cai J. Prostate cancer detection with multiparametric magnetic resonance imaging: prostate imaging reporting and data system version 1 versus version 2. *Chin Med J* 2016; **129**: 2451–9. doi: https://doi.org/10.4103/0366-6999.191771
12. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *The Lancet* 2017; **389**: 815–22. doi:

https://doi.org/10.1016/S0140-6736(16)32401-1
13. Schoots IG. MRI in early prostate cancer detection: how to manage indeterminate or equivocal PI-RADS 3 lesions? *Transl Androl Urol* 2018; **7**: 70–82. doi: https://doi.org/10.21037/tau.2017.12.31
14. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate Mr guidelines 2012. *Eur Radiol* 2012; **22**: 746–57. doi: https://doi.org/10.1007/s00330-011-2377-y
15. Vogelgesang F, Schlattmann P, Dewey M. The evaluation of bivariate mixed models in meta-analyses of diagnostic accuracy studies with SAS, Stata and R. *Methods Inf Med* 2018; **57**: 111–9. doi: https://doi.org/10.3414/ME17-01-0021
16. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**: 982–90. doi: https://doi.org/10.1016/j.jclinepi.2005.02.022
17. Alabousi M, Salameh J-P, Gusenbauer K, Samoilov L, Jafri A, Yu H, et al. Biparametric vs multiparametric prostate magnetic resonance imaging for the detection of prostate cancer in treatment-naïve patients: a diagnostic test accuracy systematic review and meta-analysis. *BJU Int* 2019; **124**: 209–20. doi: https://doi.org/10.1111/bju.14759
18. Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-Analysis of prevalence. *J Epidemiol Community Health* 2013; **67**: 974–8. doi: https://doi.org/10.1136/jech-2013-203104
19. Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. *Arch Public Health* 2014; **72**: 1–10. doi: https://doi.org/10.1186/2049-3258-72-39
20. Donato P, Morton A, Yaxley J, Teloken PE, Coughlin G, Esler R, et al. Improved detection and reduced biopsies: the effect of a multiparametric magnetic resonance imaging-based triage prostate cancer pathway in a public teaching hospital. *World J Urol* 2020; **38**: 371–9. doi: https://doi.org/10.1007/s00345-019-02774-y
21. Lee CH, Ku JY, Park WY, Lee NK, Ha HK, JY K, HK H. Comparison of the accuracy of multiparametric magnetic resonance imaging (mpMRI) results with the final pathology findings for radical prostatectomy specimens in the detection of prostate cancer. *Asia Pac J Clin Oncol* 2019; **15**: e20–7. doi: https://doi.org/10.1111/ajco.13027
22. Pal RP, Ahmad R, Trecartan S, Voss J, Ahmed S, Bazo A, et al. A single center evaluation of the diagnostic accuracy of multiparametric magnetic resonance imaging against Transperineal prostate mapping biopsy: an analysis of men with benign histology and insignificant cancer following transrectal ultrasound biopsy. *J Urol* 2018; **200**: 302–8. doi: https://doi.org/10.1016/j.juro.2018.02.072
23. Rozas G de Q, Saad LS, melo HJ de F E, Gabrielle HAA, Szejnfeld J. impact of PI-RADS V2 on indication of prostate biopsy. *Int braz j urol* 2019; **45**: 486–94.
24. Viana PCC, Horvat N, dos Santos Júnior VR, Lima TC. Romão D DOS S, Cerri LM de O, et al. *Is possible to rule out clinically significant prostate cancer using PI-RADS v2 for the assessment of prostate MRI? Int Braz J Urol* 2019; **45**: 724–31.
25. Westphalen AC, McCulloch CE, Anaokar JM, Arora S, Barashi NS, Barentsz JO, et al. Variability of the positive predictive value of PI-RADS for prostate MRI across 26 centers: experience of the Society of abdominal radiology prostate cancer Disease-focused panel. *Radiology* 2020; **296**: 76–84. doi: https://doi.org/10.1148/radiol.2020190646
26. Boesen L, Nørgaard N, Løgager V, Balslev I, Thomsen HS. Multiparametric MRI in men with clinical suspicion of prostate cancer undergoing repeat biopsy: a prospective comparison with clinical findings and histopathology. *Acta Radiol* 2018; **59**: 371–80. doi: https://doi.org/10.1177/0284185117718400
27. Lista F, Castillo E, Gimbernat H, Rodríguez-Barbero JM, Panizo J, Angulo JC. Multiparametric magnetic resonance imaging predicts the presence of prostate cancer in patients with negative prostate biopsy. *Actas Urol Esp*Ed.2015;**39**: 85–91. doi: https://doi.org/10.1016/j.acuroe.2015.01.006
28. Grey ADR, Chana MS, Popert R, Wolfe K, Liyanage SH, Acher PL. Diagnostic accuracy of magnetic resonance imaging (MRI) prostate imaging reporting and data system (PI-RADS) scoring in a transperineal prostate biopsy setting. *BJU Int* 2015; **115**: 728–35. doi: https://doi.org/10.1111/bju.12862
29. Numao N, Yoshida S, Komai Y, Ishii C, Kagawa M, Kijima T, et al. Usefulness of pre-biopsy multiparametric magnetic resonance imaging and clinical variables to reduce initial prostate biopsy in men with suspected clinically localized prostate cancer. *J Urol* 2013; **190**: 502–8. doi: https://doi.org/10.1016/j.juro.2013.02.3197
30. Daun M, Fardin S, Ushinsky A, Batra S, Nguyentat M, Lee T, et al. PI-RADS version 2 is an excellent screening tool for clinically significant prostate cancer as designated by the validated International Society of urological pathology criteria: a retrospective analysis. *Curr Probl Diagn Radiol* 2019; -**0188**: 30083–0Jun;S0363. doi: https://doi.org/10.1067/j.cpradiol.2019.06.010
31. Salami SS, Vira MA, Turkbey B, Fakhoury M, Yaskiv O, Villani R, et al. Multiparametric magnetic resonance imaging outperforms the prostate cancer prevention trial risk calculator in predicting clinically significant prostate cancer. *Cancer* 2014; **120**: 2876–82. doi: https://doi.org/10.1002/cncr.28790
32. Thompson JE, Moses D, Shnier R, Brenner P, Delprado W, Ponsky L, et al. Multiparametric magnetic resonance imaging guided diagnostic biopsy detects significant prostate cancer and could reduce unnecessary biopsies and over detection: a prospective study. *Journal of Urology* 2014; **192**: 67–74. doi: https://doi.org/10.1016/j.juro.2014.01.014
33. Gaur S, Harmon S, Mehralivand S, Bednarova S, Calio BP, Sugano D, et al. Prospective comparison of PI-RADS version 2 and qualitative in-house categorization system in detection of prostate cancer. *J Magn Reson Imaging* 2018; **48**: 1326–35. doi: https://doi.org/10.1002/jmri.26025
34. Hsieh P-F, Li W-J, Lin W-C, Chang H, Chang C-H, Huang C-P, et al. Combining prostate health index and multiparametric magnetic resonance imaging in the diagnosis of clinically significant prostate cancer in an Asian population. *World J Urol* 2020; **38**: 1207–14. doi: https://doi.org/10.1007/s00345-019-02889-2
35. Luzzago S, Musi G, Catellani M, Russo A, Di Trapani E, Mistretta FA, et al. Multiparametric magnetic-resonance to confirm eligibility to an active surveillance program for low-risk prostate cancer: intermediate time results of a third referral high volume centre active surveillance protocol. *Urol Int* 2018; **101**: 56–64. doi: https://doi.org/10.1159/000488772
36. Zhang K, Chen R, Alberts AR, Zhu G, Sun Y, Roobol MJ. Distribution of prostate imaging reporting and data system score and diagnostic accuracy of magnetic resonance imaging-targeted biopsy: comparison of an Asian and European cohort. *Prostate Int* 2019; **7**: 96–101. doi: https://doi.org/10.1016/j.prnil.2018.10.001
37. Jordan EJ, Fiske C, Zagoria RJ, Westphalen AC. Evaluating the performance of PI-RADS V2 in the non-academic setting. *Abdom Radiol* 2017; **42**: 2725–31. doi: https://doi.org/10.1007/s00261-017-1169-5

38. Thompson JE, van Leeuwen PJ, Moses D, Shnier R, Brenner P, Delprado W, et al. The diagnostic performance of multiparametric magnetic resonance imaging to detect significant prostate cancer. *J Urol* 2016; **195**: 1428–35. doi: https://doi.org/10.1016/j.juro.2015.10.140
39. Zhao C, Gao G, Fang D, Li F, Yang X, Wang H, et al. The efficiency of multiparametric magnetic resonance imaging (mpMRI) using PI-RADS version 2 in the diagnosis of clinically significant prostate cancer. *Clin Imaging* 2016; **40**: 885–8. doi: https://doi.org/10.1016/j.clinimag.2016.04.010
40. Osses DF, van Asten JJ, Kieft GJ, Tijsterman JD. Prostate cancer detection rates of magnetic resonance imaging-guided prostate biopsy related to prostate imaging reporting and data system score. *World J Urol* 2017; **35**: 207–12. doi: https://doi.org/10.1007/s00345-016-1874-7
41. Greer M, Shih J, Lay N, Barrett T, Radiology LKB-. Undefined. validation of the dominant sequence paradigm and role of dynamic contrast-enhanced imaging in Pi-raDs version 2 1. *Natl Institutes Heal 10 n Radiol* 2017; **285**: 859–692017 Dec.
42. Mussi TC, Martins T, Garcia RG, Filippi RZ, Lemos GC, Baroni RH. Are dynamic contrast-enhanced images necessary for prostate cancer detection on multiparametric magnetic resonance imaging? *Clin Genitourin Cancer* 2017; **15**: e447–54. doi: https://doi.org/10.1016/j.clgc.2016.10.001
43. Choi MH, Kim CK, Lee YJ, Jung SE. Prebiopsy biparametric MRI for clinically significant prostate cancer detection with PI-RADS version 2: a multicenter study. *AJR Am J Roentgenol* 2019; **212**: 839–46. doi: https://doi.org/10.2214/AJR.18.20498