Urologic Oncology: Seminars and Original Investigations 37 (2019) 299.e1–299.e6

UROLOGIC ONCOLOGY

Difference in MRI-guided biopsy cancer detection rates between individual clinicians

Niranjan J. Sathianathen, M.B.B.S.[a,*], Christopher A. Warlick, M.D., Ph.D.[a], Ayman Soubra, M.D.[a], Priyamvadha Balaji[a], Gregory J. Metzger, M.D., Ph.D.[b], Benjamin Spilseth, M.D.[b], Paari Murugan, M.D.[c], Maria Ordonez, M.D.[a], Christopher J. Weight, M.D.[a], Badrinath R. Konety, M.D., M.B.A.[a]

[a] *Department of Urology, University of Minnesota, Minneapolis, MN*
[b] *Department of Radiology, University of Minnesota, Minneapolis, MN*
[c] *Department of Pathology, University of Minnesota, Minneapolis, MN*

Received 8 August 2018; received in revised form 13 January 2019; accepted 31 January 2019

Abstract

Objective: A number of studies have described the overall institutional learning curve for magnetic resonance imaging-guided biopsy but none have evaluated differences and interactions between clinicians. Therefore, we aim to measure and compare the cancer detection rates between individual radiologists and urologists at a single academic institution.

Methods: A consecutive sample of patients undergoing magnetic resonance imaging-guided biopsy at a single institution were included for analysis. The detection of any and clinically significant (Gleason score ≥3+4) prostate cancer was compared between radiologists and urologists after adjusting for relevant demographic and clinical characteristics. Analysis was conducted on a perlesion basis and only the results of the targeted cores were considered in the primary analysis.

Results: Two hundred eighty-one patients with 418 lesions were included in the study. Prostate cancer of any grade was detected in 43.7% (183/418) of targeted lesions. There was no difference in the distribution of Prostate Imaging Reporting and Data System (PIRADS) scores attributed by each radiologist ($p = 0.43$). The individual radiologist cancer detection rate for both overall and clinically significant cancer was similar across each PIRADS score except for the detection of any cancer in PIRADS 3 lesions ($p = 0.03$). There was no difference in the detection rates of any grade or clinically significant cancer between urologists.

Conclusion: This single institutional analysis found that the performance of radiologists and urologists was largely comparable. The only variation observed was among radiologists for PIRADS 3 lesions.

Keywords: Image-guided biopsy; Magnetic resonance imaging; Prostate cancer

1. Introduction

Multiparametric magnetic resonance imaging (mpMRI) has come to the forefront of prostate imaging following an abundance of literature demonstrating its superiority over other modalities in identifying suspicious lesions. Using a reference of transperineal mapping biopsy, the PROMIS study reported that mpMRI had a sensitivity of 93% for clinically significant disease, compared to 48% for transrectal ultrasound-guided biopsy [1]. Following delineation of regions of interest (ROI), MRI-guided biopsy (MRGB) has enabled targeting of these regions to improve diagnostic performance. Compared to 'blind' biopsy methods, MRGB has been reported to diagnose 30% more high-risk cancer and decrease the diagnosis of indolent disease by 17% [2,3].

A number of factors have been suggested to influence biopsy yield. Patient factors such as prostate-specific

Funding: NJS has received support from the Cloverfields Foundation and The Institute for Prostate and Urologic Cancers (University of Minnesota).

*Corresponding author. Tel.: +1-660-334-0774; fax: +1-612-626-0428.
E-mail address: nsathian@umn.edu (N.J. Sathianathen).

https://doi.org/10.1016/j.urolonc.2019.01.032
1078-1439/

antigen (PSA) level and prostate volume have been demonstrated to influence the likelihood of detecting cancer on MRGB [4]. The mode of image registration has been shown to impact cancer detection with a higher detection rate in software registration than with cognitive registration [5]. The true effect of this is not clear as there have been conflicting findings in the literature with a review finding that clinically significant cancer detection was comparable across MRI target biopsy techniques [6]. Furthermore, as is to be expected with new technologies, institutional experience has been shown to be an important factor in diagnostic performance. At 1 institution, over an 18-month period, the MRGB cancer detection increased significantly from 27% to 63% [7]. Similarly, Meng and colleagues reported that their clinically significant cancer detection rate for PIRADS 4/5 lesions increased by 26% over a 4-year period [8]. Yet it is not completely clear whether a difference exists at an individual clinician level and the nature of any interaction between clinicians. Hence, this study aims to assess whether there is a difference in targeted biopsy prostate cancer detection rate between individual radiologists and urologists.

2. Methods

2.1. Patients

A consecutive sample of men from July 2014 to June 2017 undergoing MRGB using UroNav (Invivo, Gainesville, FL) ultrasound fusion software at a single academic institution were included in this retrospective study. This was a heterogeneous cohort consisting of men undergoing biopsy in 1 of 3 clinical settings: primary biopsy, previous negative biopsy, or active surveillance follow-up.

2.2. Imaging and biopsy protocol

Diagnostic mpMRI was performed using a 3-Tesla Siemens scanner using an endorectal (BPX-30; Medrad, Indianola, PA) or pelvic phased array coil (Magnetom Trio-Tim or Skyra; Siemens Healthcare; Erlangen, Germany). The detailed MRI protocol and acquisition parameters have been previously described [9]. All scans were interpreted by 1 of 3 fellowship-trained radiologists in accordance with the PIRADS version used at the time. Because this study aimed to reflect real-world clinical practice, older prostate MRI images that were reported using PIRADS v1, were not retrospectively re-evaluated using the updated PIRADS v2 system. Radiologists 1, 2, and 3 had 3, 10, and 1 years of experience interpreting prostate MRI, respectively. MRIs were interpreted as part of routine clinical practice and radiologists were not blinded to the patient's clinical information at the time of interpretation.

At the discretion of the treating urologist, patients proceeded to undergo a manual MRGB with at least 2 cores obtained from each identified lesion followed by a 12–14 core systematic biopsy. The general practice was to biopsy men with a PIRADS 3–5 lesion; PIRADS 2 lesions were also commonly biopsied if there was an index lesion of higher grade especially during the early phase of our MRI program [10]. All these lesions were included for analysis. Cases with a negative mpMRI were not included in this study because they were not subjected to a targeted biopsy. Urologists 1, 2, 3, and 4 have been in independent clinical practice for 14, 8, 6, and 4 years, respectively. Urologists 1 and 2 had been performing in-bore direct MRI-targeted biopsies prior to the commencement of this study. No urologists had prior ultrasound fusion experience prior to the cases reported in this study. Radiologists were not involved during the biopsy procedure.

2.3. Statistical analysis

The primary outcome was overall and clinically significant cancer detection rate (CDR) by targeted biopsy and by individual radiologists and urologists. Clinically significant cancer was defined as Gleason score $3+4=7$ (grade group2) or greater. Analysis was performed on a perlesion basis using the overall reported PIRADS score which combined scores from each of the imaging series (T2 weighted, diffusion weighted, and dynamic contrast enhanced). The results of systematic biopsy were not considered in the primary analyses of this study. Cancer detection rate between individual clinicians was compared using a chi-square test and multivariable logistic regression models were used to adjust for relevant clinical and demographic variables (age, prostate-specific antigen (PSA), clinical setting (primary biopsy, previous negative biopsy, active surveillance), mpMRI prostate volume, PIRADS score). A chi-square trend test was used to evaluate a learning curve. Operators that had been involved in interpreting mpMRIs or performing biopsy in less than 5% of the sample were excluded. All p values were 2 sided and statistical significance was set at the 0.05 level. Data analysis was performed in R (R Foundation for Statistical Computing, Vienna, Austria) version 3.4.

3. Results

Overall, 281 men proceeded to a MRGB. Complete demographic data on the included cohort is outlined in Table 1. Among the included cohort, there were 418 discrete lesions of interest targeted on biopsy. There was a difference between the proportion of patients managed by each radiologist and urologist combination (Supplementary Table 1).

Prostate cancer of any grade was detected in 43.7% (183/418) of targeted lesions. The clinically significant CDR was 23.4% (n = 98/418). There was a significant association between PIRADS score and CDR with cancer of any grade being detected in 6.9% of PIRADS 2 lesions, 27.4% of PIRADS 3 lesions, 46.6% of PIRADS 4 lesions,

Table 1
Patient demographics

n	281
Median age, years (IQR)	65.2 (60.1–70.3)
Median PSA, ng/mL(IQR)	7.6 (5.1–10.7)
Clinical setting, n (%)	
Primary biopsy	113 (40.2)
Previous negative biopsy	76 (27.0)
Active surveillance	91 (32.4)
Median mpMRI prostate volume, mL (IQR)	43.9 (31.4–62.7)
Median number of lesions (IQR)	1 (1-2)
Median number of targeted cores per lesion (IQR)	2 (2-2)
Number of mpMRIs interpreted, n (%)	
Radiologist 1	134 (47.7)
Radiologist 2	129 (45.6)
Radiologist 3	19 (6.8)
Number of MRGB performed, n (%)	
Urologist 1	104 (37.0)
Urologist 2	93 (33.1)
Urologist 3	44 (15.7)
Urologist 4	40 (14.2)

and 82.8% of PIRADS 5 lesions ($p < 0.01$). The corresponding clinically significant cancer detection rate for each PIRADS score was 0%, 7.7%, 24.3%, and 60.9%, respectively ($p < 0.01$). The PIRADS version used at the time of scan interpretation had no effect on the CDR of any grade ($p = 0.09$) or significant ($p = 0.14$) cancers.

Radiologists 1, 2, and 3 interpreted 134 (47.7%), 129 (45.6%), and 19 (6.8%) scans, respectively. There was no significant difference in the distribution of PIRADS scores by radiologist ($p = 0.43$), as depicted in Fig. 1. The proportion of lesions graded by each radiologist as PIRADS 3, 4, and 5 ranged from 14.8% to 30.9%, 46.6% to 63.0%, and 13.1% to 18.5%, respectively. The individual radiologist CDR for both overall and clinically significant cancer was similar across each PIRADS score (Table 2) except for the detection of any cancer in PIRADS 3 lesions ($p = 0.025$). On sensitivity analysis (when excluding radiologist 3 who only interpreted 19 scans in this cohort), the difference between radiologist 1 and 2 in the detection of any cancer in PIRADS 3 lesions was still significant.

Biopsy was performed by 4 urologists who respectively did 104 (37.0%), 93 (33.1%), 42 (15.7%) and 40 (14.2%) procedures. There was no significant difference in biopsy indication between urologists (Supplementary Table 2). There was no significant difference in the cancer detection rate for clinically significant or cancer of any grade between urologists.

On multivariate regression analysis there was no difference among radiologists nor urologists in the diagnosis of any or clinically significant prostate cancer. PSA, prostate volume and PIRADS score acted as predictors of cancer detection. The version of PIRADS score used to interpret the scans had no effect.

When also including the results of the systematic biopsy, there was only a difference in clinically significant cancer detection rate favoring urologist 2 compared to urologist 1 [OR 2.6, 95%CI 1.2–5.7]. There was no difference between radiologists or other urologists in the diagnosis of clinically significant or cancer of any grade. The referent clinician was also changed in subsequent iterations of the model and this finding remained consistent (data not shown).

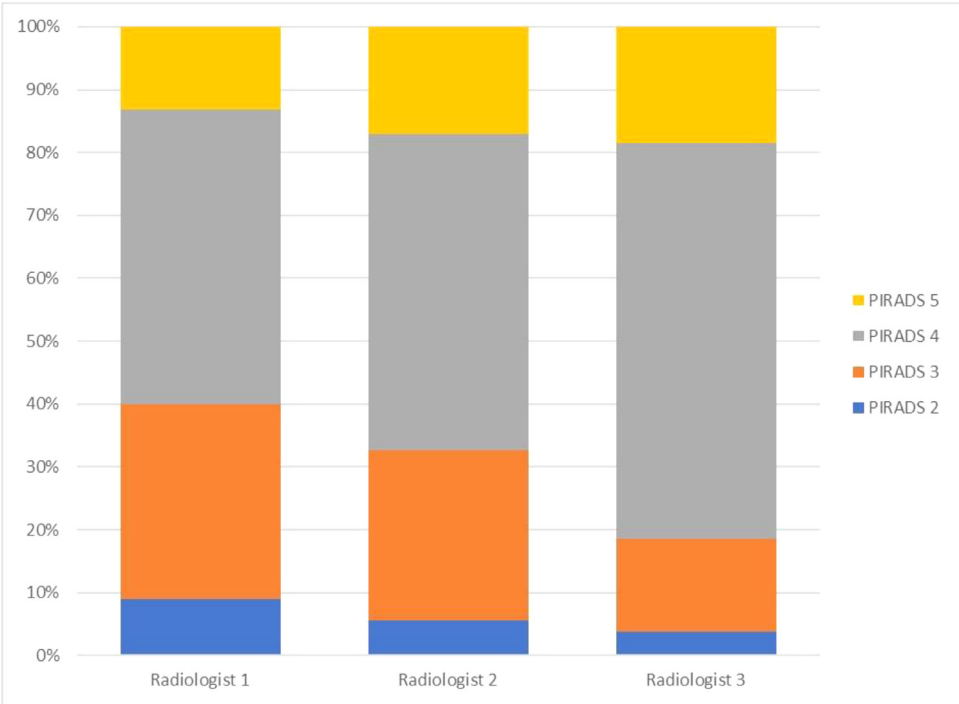

Fig. 1. Distribution of PIRADS scores by radiologists.

Table 2
Cancer detection rates on targeted biopsy by individual radiologists and urologists

	Radiologist 1	Radiologist 2	Radiologist 3		p value
PIRADS 3					
Any cancer, % (n)	18.6 (12)	38.9 (21)	0.0 (0)		0.025
Significant cancer, % (n)	5.1 (3)	11.1 (6)	0.0 (0)		0.409
PIRADS 4					
Any cancer, % (n)	46.1 (44)	44.0 (46)	64.7 (11)		0.284
Significant cancer, % (n)	22.5 (20)	25.0 (25)	29.4 (5)		0.807
PIRADS 5					
Any cancer, % (n)	92.0 (23)	76.5 (28)	80.0 (4)		0.291
Significant cancer, % (n)	68.0 (17)	61.8 (23)	20.0 (1)		0.132
	Urologist 1	Urologist 2	Urologist 3	Urologist 4	p value
PIRADS 3					
Any cancer, % (n)	26.0 (14)	42.4 (14)	23.1 (3)	9.5 (2)	0.063
Significant cancer, % (n)	6.0 (3)	12.1 (4)	7.7 (1)	4.8 (1)	0.713
PIRADS 4					
Any cancer, % (n)	47.1 (37)	50.7 (36)	45.2 (14)	38.2 (14)	0.689
Significant cancer, % (n)	21.4 (15)	29.6 (21)	22.6 (7)	20.6 (7)	0.638
PIRADS 5					
Any cancer, % (n)	78.9 (16)	81.5 (23)	85.7 (12)	100.0 (4)	0.766
Significant cancer, % (n)	57.9 (12)	59.3 (17)	64.3 (9)	75.0 (3)	0.918

In an attempt to characterize a learning curve, analysis was performed in consecutive 10 case increments. The yield of significant cancer for urologists and radiologists are shown in Fig. 2. The Cochrane Armitage test for trend was significant for all the clinicians analyzed, except for urologist 3.

4. Discussion

This study found that the MRGB cancer detection rate by individual radiologists and urologists are largely comparable. Radiologists only differed in the yield of any cancer in PIRADS 3 lesions and did not differ in yield for clinically significant cancer for any PIRADS score. There was no difference among urologists in detecting prostate cancer using MRGB. Furthermore, CDRs of individual clinicians were also largely comparable when the results of the systematic biopsy were also considered. These radiologist results are a contrast to the experience reported by Sonn and colleagues where individual radiologists had an influence on the diagnostic performance [11]. However, there were a greater number of radiologists (9) who interpreted prostate mpMRIs in the aforementioned study and therefore there is a greater probability of observing a statistically significant difference between individuals because of the increased number of comparisons. Furthermore, no radiologist in Sonn et al. read more than 70 scans and 2 read fewer than 20 scans so it is possible that the differences observed are due to individuals being in various stages of the learning curve in their study. Moreover, it is not surprising that an inconsistency in CDRs was seen among PIRADS 3 lesions which are “equivocal” and introduce uncertainty to the clinical setting [12]. Although the explanation for this is likely multifactorial, heterogenous imaging characteristics of equivocal lesions may be an important driver and has led to suggestions of further subclassification or adjustments to the PIRADS system [12,13]. The discrepancy in these findings highlights the importance of individual clinicians, not just institutions, auditing their own performance so that any potential areas for improvement can be identified.

Although the performance of clinicians was mostly similar, there were some intricate differences identified in CDRs between individuals. Studies have reported that institutional experience influences diagnostic performance of MRGB with an increase in CDR from 42% to 81% over a time period of 18 months [7]. There are many factors other than individual clinician performance that may play a role in institutional differences including technological differences, MRI protocols and image quality, pathologic interpretation, as well as shifting indications for biopsy and resultant shifting likelihood of malignancy. Clinician experience may explain the lower absolute CDR which approached statistical significance in PIRADS 5 lesions because radiologist 3 who had interpreted the least scans and is still likely in the learning curve. However, it is unlikely that operator experience is the only reason for the difference seen in CDR; rather patient factors are likely underpinning the observed differences in biopsy yield between clinicians. Of note, the lowest CDR was seen in the urologist with the most clinical experience in terms of number of fusion biopsies performed, suggesting that the learning curve for guided biopsy is not accelerated by greater clinical experience. Similarly, there was a discrepancy observed in the diagnosis of any grade cancer in PIRADS 3 lesions between the 2 most experienced radiologists. There are many possible pitfalls that may lead to “missing” lesions on biopsy including inaccurate fusion or tracking of targeted biopsies. It is possible that differences

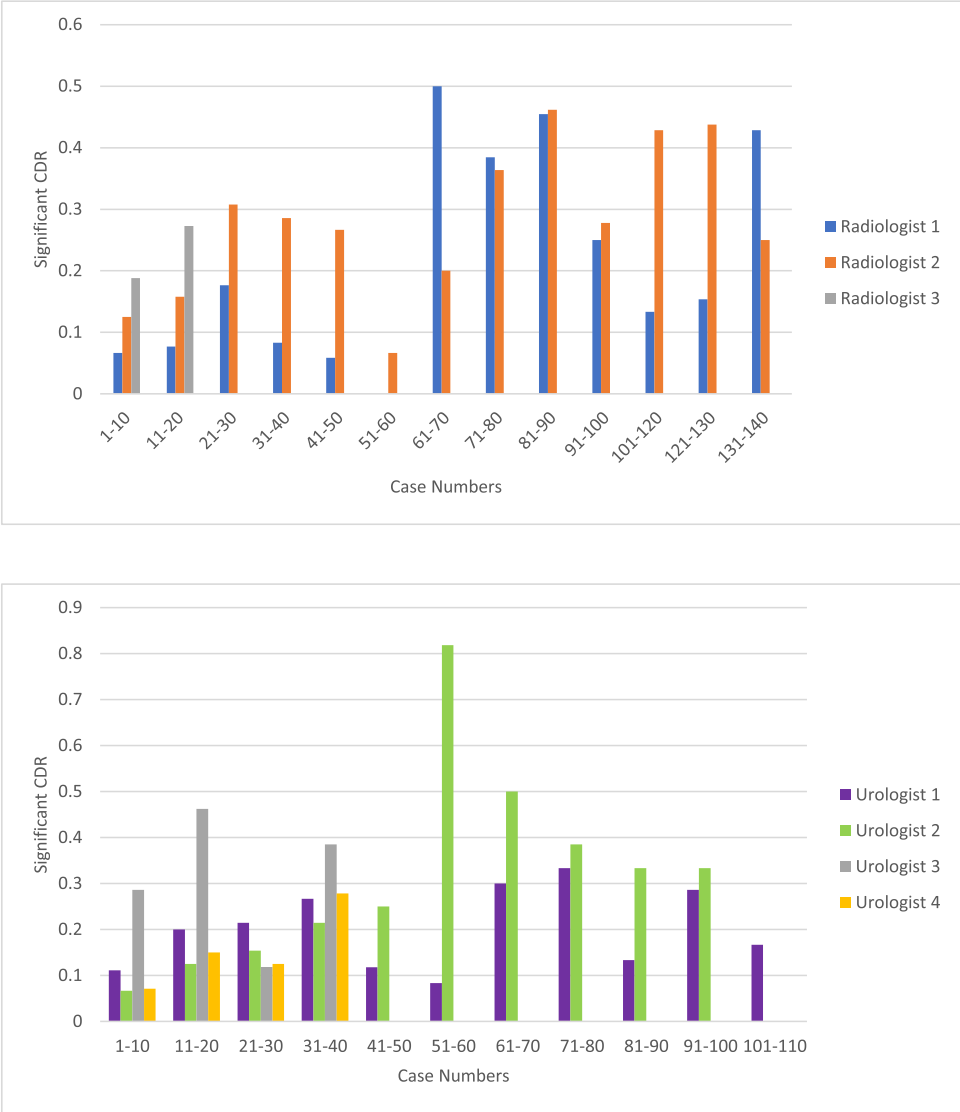

Fig. 2. Yield by case numbers for highest volume radiologists and urologists.

in technique, such as if cores were obtained from both sagittal and axial planes rather than only 1, could be partly contributing to the observed relationship [14]. The number of targeted cores has also been shown to affect biopsy accuracy [15], but there was no difference between urologists in the numbers of targeted cores obtained ($p = 0.81$, data not shown). It is also possible that additional differences in patient characteristics that we could not adjust for in our analysis, for example, lesion size and location, are underlying the differences in CDR [16]. This explanation is supported by the statistically significant univariate results which generally did not achieve statistical significance in multivariable models. Nonetheless, in light of the recently published PRECISION randomized trial which demonstrated that targeted biopsy only was noninferior to systematic biopsy [3], information on individual clinician biopsy yield can also inform decision making on whether it is safe to omit systematic sampling.

Primarily due to the pragmatic design of this study which was intended to reflect real-world clinical practice, there are important limitations that should be considered when interpreting these results. Although there were general guidelines, it was not possible to completely account for selection bias in terms of whom urologists chose to biopsy as a multitude of factors are involved in this decision. Additionally, prostate mpMRIs were only interpreted by a single radiologist and scans which were reported in accordance to PIRADS v1 were not retrospectively reinterpreted because the biopsy was performed using the report at the time and it was the aim of this study to report on outcomes during routine clinical practice. However, our results show that the version of PIRADS score used at the time of MRI interpretation had no impact on CDRs. Furthermore, there is a complex interaction between radiologists' experience and urologists' outcomes that we have probably not completely adjusted for, but it could be reasonably

suggested that because all radiologists were fellowship trained that they had passed the learning curve prior to inception of this study. Additionally, it is challenging to determine the effect of previous non-MRGB experience on the learning curve of MRGB. Image registration was performed using the UroNav system and therefore it is possible that differences in CDR between individuals are amplified when using cognitive registration although it has been reported that outcomes are comparable between the techniques [6,17]. Moreover, this study was performed in the academic setting and therefore its results may not be generalizable to nonacademic institutions [18].

5. Conclusion

We found that the MRGB cancer detection rate between radiologists and urologists to be largely comparable. However, we did observe some intricate differences on secondary analysis, which highlights the importance of individual clinicians auditing their own performance.

Conflicts of interest

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.urolonc.2019.01.032.

References

[1] 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. Lancet (London, England) 2017;389:815–22.

[2] Siddiqui MM, Rais-Bahrami S, Turkbey B, George AK, Rothwax J, Shakir N, et al. Comparison of MR/ultrasound fusion−guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA 2015;313:390–7.

[3] Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. Engl J Med 2018;378:1767–77.

[4] Hansen NL, Kesch C, Barrett T, Koo B, Radtke JP, Bonekamp D, et al. Multicentre evaluation of targeted and systematic biopsies using magnetic resonance and ultrasound image-fusion guided transperineal prostate biopsy in patients with a previous negative biopsy. BJU Int 2017;120:631–63.

[5] Oberlin DT, Casalino DD, Miller FH, Matulewicz RS, Perry KT, Nadler RB, et al. Diagnostic value of guided biopsies: fusion and cognitive-registration magnetic resonance imaging versus conventional ultrasound biopsy of the prostate. Urology 2016;92:75–9.

[6] Wegelin O, van Melick HHE, Hooft L, Bosch J, Reitsma HB, Barentsz JO, et al. Comparing three different techniques for magnetic resonance imaging-targeted prostate biopsies: a systematic review of in-bore versus magnetic resonance imaging-transrectal ultrasound fusion versus cognitive registration. is there a preferred technique? Eur Urol 2017;71:517–31.

[7] Gaziev G, Wadhwa K, Barrett T, Koo BC, Gallagher FA, Serrao E, et al. Defining the learning curve for multiparametric magnetic resonance imaging (MRI) of the prostate using MRI-transrectal ultrasonography (TRUS) fusion-guided transperineal prostate biopsies as a validation tool. BJU Int 2016;117:80–6.

[8] Meng X, Rosenkrantz AB, Huang R, Deng FM, Wysock JS, Bjurlin M, et al. The institutional learning curve for MRI-US fusion-targeted prostate biopsy: temporal improvements in cancer detection over four years. J Urol 2018;200:1022–9.

[9] Metzger GJ, Kalavagunta C, Spilseth B, Bolan PJ, Li X, Hutter D, et al. Detection of prostate cancer: quantitative multiparametric MR imaging models developed using registered correlative histopathology. Radiology 2016;279:805–16.

[10] Rosenkrantz AB, Verma S, Choyke P, Eberhardt SC, Eggener SE, Gaitonde K, et al. Prostate magnetic resonance imaging and magnetic resonance imaging targeted biopsy in patients with a prior negative biopsy: a consensus statement by AUA and SAR. J Urol 2016;196:1613–8.

[11] Sonn GA, Fan RE, Ghanouni P, Wang NN, Brooks JD, Loening AM, et al. Prostate magnetic resonance imaging interpretation varies substantially across radiologists. Eur Urol Focus 2017:pii: S2405-4569(17)30266-3.

[12] Hansen NL, Koo BC, Warren AY, Kastner C, Barrett T. Sub-differentiating equivocal PI-RADS-3 lesions in multiparametric magnetic resonance imaging of the prostate to improve cancer detection. Eur J Radiol 2017;95:307–13.

[13] Rosenkrantz AB, Babb JS, Taneja SS, Ream JM. Proposed adjustments to PI-RADS version 2 decision rules: impact on prostate cancer detection. Radiology 2017;283:119–29.

[14] Hong CW, Rais-Bahrami S, Walton-Diaz A, Shakir N, Su D, George AK, et al. Comparison of magnetic resonance imaging and ultrasound (MRI-US) fusion-guided prostate biopsies obtained from axial and sagittal approaches. BJU Int 2015;115:772–9.

[15] Calio BP, Sidana A, Sugano D, Gaur S, Maruf M, Jain AL, et al. Risk of upgrading from prostate biopsy to radical prostatectomy pathology — Does saturation biopsy of index lesion during multiparametric magnetic resonance imaging-transrectal ultrasound fusion biopsy help? J Urol 2018;199:976–82.

[16] Boesen L, Norgaard N, Logager V, Balslev I, Thomsen HS. Where do transrectal ultrasound- and magnetic resonance imaging-guided biopsies miss significant prostate cancer? Urology 2017;110:154–60.

[17] Arsov C, Rabenalt R, Blondin D, Quentin M, Hiester A, Godehardt E, et al. Prospective randomized trial comparing magnetic resonance imaging (MRI)-guided in-bore biopsy to MRI-ultrasound fusion and transrectal ultrasound-guided prostate biopsy in patients with prior negative biopsies. Eur Urol 2015;68:713–20.

[18] Hansen NL, Koo BC, Gallagher FA, Warren AY, Doble A, Gnanapragasam V, et al. Comparison of initial and tertiary centre second opinion reads of multiparametric magnetic resonance imaging of the prostate prior to repeat biopsy. Eur Radiol 2017;27:2259–66.