FULL PAPER

Comparison of Likert and PI-RADS version 2 MRI scoring systems for the detection of clinically significant prostate cancer

[1]JERIES P ZAWAIDEH, [1,2]EVIS SALA, [1]MARIA PANTELIDOU, [1,2]NADEEM SHAIDA, [1]BRENDAN KOO, [1]IZTOK CAGLIC, [2,3]ANNE Y WARREN, [4]LUCA CARMISCIANO, [2,5]KASRA SAEB-PARSY, [2,5]VINCENT J GNANAPRAGASAM, [2,5]CHRISTOF KASTNER and [1,2]TRISTAN BARRETT

[1]Department of Radiology, Addenbrooke's Hospital and University of Cambridge, Cambridge, UK
[2]CamPARI Prostate Cancer Group, Addenbrooke's Hospital and University of Cambridge, Cambridge, UK
[3]Department of Pathology, Addenbrooke's Hospital, Cambridge, UK
[4]Department of health Sciences (DISSAL), Biostatistics section, University of Genoa, Genoa, Italy
[5]Department of Urology, Addenbrooke's Hospital, Cambridge, UK

Address correspondence to: MD Jeries P Zawaideh
E-mail: *jeriespz89@gmail.com*

Objective: To compare the performance of Likert and Prostate Imaging–Reporting and Data System (PI-RADS) multiparametric (mp) MRI scoring systems for detecting clinically significant prostate cancer (csPCa).

Methods: 199 biopsy-naïve males undergoing prostate mpMRI were prospectively scored with Likert and PI-RADS systems by four experienced radiologists. A binary cut-off (threshold score ≥3) was used to analyze histological results by three groups: negative, insignificant disease (Gleason 3 + 3; iPCa), and csPCa (Gleason ≥3 +4). Lesion-level results and prostate zonal location were also compared.

Results: 129/199 (64.8%) males underwent biopsy, 96 with Likert or PI-RADS score ≥3, and 21 with negative MRI. A further 12 patients were biopsied during follow-up (mean 507 days). Prostate cancer was diagnosed in 87/199 (43.7%) patients, 65 with (33.6%) csPCa. 30/92 (32.6%) patients with negative MRI were biopsied, with an NPV of 83.3% for cancer and 86.7% for csPCa.

Likert and PI-RADS score differences were observed in 92 patients (46.2%), but only for 16 patients (8%) at threshold score ≥3. Likert scoring had higher specificity than PI-RADS (0.77 *vs* 0.66), higher area under the curve (0.92 *vs* 0.87, *p* = 0.002) and higher PPV (0.66 *vs* 0.58); NPV and sensitivity were the same. Likert had more five score results (58%) compared to PI-RADS (36%), but with similar csCPa detection (81.0 and 80.6% respectively). Likert demonstrated lower proportion of false positive in the predominately AFMS-involving lesions.

Conclusion: Likert and PI-RADS systems both demonstrate high cancer detection rates. Likert scoring had a higher AUC with moderately higher specificity and lower positive call rate and could potentially help to reduce the number of unnecessary biopsies performed.

Advances in knowledge: This paper illustrates that the Likert scoring system has potential to help urologists reduce the number of prostate biopsies performed.

INTRODUCTION

International guidelines now recommend multiparametric (mp) MRI as the initial diagnostic test in males referred with a suspicion of prostate cancer.[1,2] The Prostate Imaging–Reporting and Data System (PI-RADS) guidelines initially developed in 2012 with further updates in 2015 and 2019[3–5] aim to improve MR image acquisition and quality, and also to standardize the interpretation of prostate MRI. Likert-based scoring systems pre-date the development of PI-RADS and are currently advocated in UK guidelines for mpMRI interpretation.[6–8] Key differences include the incorporation of clinical information into

Likert assessment and an equal evaluation of all mpMRI sequences, in contrast to the zone-specific dominant sequence paradigm of PI-RADS.[9]

There have been a limited number of studies comparing the two scoring systems, mainly involving the first version of PI-RADS, which pre-dates the introduction of the dominant sequence model.[9–11] A more recent study evaluated Likert with PI-RADS v. 2, however, this only involved 1.5 T imaging, and imaging included a mismatch between T2 and diffusion-weighted imaging, against current recommendations.[12] Furthermore, a UK consensus panel has

previously highlighted the need for further evidence comparing the two systems.[6] The purpose of this study was therefore to directly compare the diagnostic performance of the Likert and PI-RADS v. 2 systems for the detection of all cancer and clinically significant prostate cancer (csPCa).

METHODS

Consecutive biopsy-naïve patients undergoing 3 T prostate mp-MRI in a university hospital from May 2018 to October 2018 were included.. The local ethics committee approved this research project as a service evaluation and waived the need for informed consent (**CUH/018/PRN7917**). Patients were referred from primary care with a suspicion of prostate cancer based on a raised prostate-specific antigen (PSA) and/or positive digital rectal examination. Patients were prospectively scored using both the Likert and the PI-RADS scoring systems. Exclusion criteria included patients on active surveillance (n = 13), prior treatment for prostate cancer (2) significant motion artefact (9), or presence of hip replacement (n = 2). Following this, 199 patients were included for analysis.

MRI parameters

Prostate MRI was performed on a 3 T HDx Discovery MR750 HDx (GE Healthcare, Waukesha, USA) using a 32-channel phased array coil. Axial, sagittal and coronal planes T_2 weighted FR-FSE images of the prostate were acquired [echo time/repetition time (TE/TR) = 85/3700 ms; field of view (FOV) 24 × 24 cm; matrix 256 × 256; slice thickness 3 mm]; DWI was performed using a SE pulse sequence (TE/TR = 60/3400 ms; matrix 256 × 256; slice thickness 3 mm). The following b-values were acquired: b-150, b-750, b-1,000, b-1,400 and b-2,000 s/mm^2; apparent diffusion coefficient maps were automatically calculated. DCE-MRI was performed as an axial 3D-FSPGR sequence (TR/TE 4.088/1.788 ms, FOV 24 × 24 cm) following bolus injection of Gadobutrol (Gadovist, Bayer Healthcare) at 3 ml s^{-1} (dose 0.1 mmol/kg), with a temporal resolution of 7 s.

Image analysis

Mp-MRIs were prospectively reported by one of four experienced uro-radiologists. Readers had full access to all relevant clinical information, including PSA, findings on digital rectal examination, and any othermedical history. Scoring was performed according to PI-RADS v. 2,[3,13] and using a Likert system, which is the default system used in our department, as previously described,[14] wherein 1: highly unlikely to have clinically significant prostate cancer (csPCa), 2: csPCaunlikely, 3: indeterminate for csPCa, 4: csPCalikely, 5: highly likely to have csPCa; Supplementary Table 1. Both Likert and PI-RADS scores were reported simultaneously at the time of initial review. All scores ≥ 3 were reviewed in a multidisciplinary team setting to determine whether targeted biopsy should be performed. In cases of a negative MRI, a decision to perform a systematic biopsy was based on clinical risk factors, including family history of prostate cancer, absolute PSA ≥10 ng ml^{-1} and PSA-density ≥0.15 ng/ml/cm^3. Any patient not undergoing a biopsy was followed-up for at least 14 months, with at least one subsequent PSA reading.

Biopsy technique

When a target lesion was identified, biopsy was performed using a MRI/ultrasound fusion technique by either a transrectal (UroNav^{TM}; InVivo Corp) or transperineal approach (Biopsee^{TM}; Medcom). At least two biopsy cores were taken from each target lesion prior to systematic biopsy cores. A genitourinary pathologist evaluated all biopsies using the Gleason score according to International Society of Urological pathology (ISUP) 2014 recommendations; all the results were confirmed in a multidisciplinary team setting.[15] The final histology result was used as reference standard for outcome purposes.

Statistical analysis

Statistical analyses were conducted using "R" v. 3.5.3 (2019; The R Foundation for Statistical Computing). General characteristics with mean and standard deviation (SD) for normally distributed continued variables, and median and interquartile ranges (IQRs) for skewed continued variables were calculated. For cases with multiple lesions, only the dominant "index lesion" was assessed. Outcomes were clustered into three clinically meaningful groups based on pathology results: negative, Gleason 3 + 3 (insignificant prostate cancer, iPCa), and Gleason ≥ 3+4 (csPCa) and reported as counts and percentages. A radiological cut-off of score ≥3 was applied for both Likert and PI-RADS scores based on the stated criteria for each reporting tool, outcomes were reported by sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), absolute false positive (FP) and negative (FN) counts, receiver operating characteristic (ROC) curves and compared with Delong test to evaluate statistically significant differences. Using the binary clustered results, McNemar's test was used to detect significant changes in MRI score between the two reporting score systems. Score outcome differences between the two methods were separately analyzed and compared to biopsy outcome. For statistical analysis, target lesions were stratified according to prostate zonal involvement (PZ and TZ), lesions arising from the TZ but involving mostly the anterior fibromuscular stroma (AFMS) were separately assessed;absolute true positive (TP) and FP and FN counts were calculated for csPCa and all PCa; to mitigate the biased estimation due unavailable biopsy results (MRI TN and MRI FNs) PPV and false discovery rate (FDR) were obtained for the two scores.

RESULTS

199 patients were assessed, with a mean age of 64 years (range 44–84) and median PSA of 6.00 (IQR 4.38–8.74); Table 1.

Pathological outcomes

In total, 117/199 (58.7%) patients underwent initial biopsy, including 96 with a positive MRI (Likert/PI-RADS score ≥3) and 21 patients with a negative MRI. A further 12 patients were biopsied during the follow-up period, thus in total 129 (64.8%) patients were biopsied, of which 9 patients were negative on both Likert and PI-RADS scoring. Prostate cancer was diagnosed in 87/199 (43.7%) patients, including 65/199 (33.6%) clinically significant tumors, Gleason ≥ 3+4. The 70 patients with no biopsy were followed-up clinically for an average of 16 months (median 16, range 14–18 months). In 92 patients, the final MRI score was ≤2, 30/92 (32.6%) of those patients were biopsied, with

Table 1. General characteristics of included patients

	Overall	Negative	Gleason 3 + 3	Gleason ≥ 3+4
Number	199	112	22	65
Age [years] Mean (SD)	64.0 (8.4)	61.9 (8.0)	63.3 (11.3)	67.8 (6.4)
PSA level [ng/ml] Median [IQR]	6.00 [4.38, 8.74]	5.65 [4.21–8.11]	4.85 [3.87–6.63]	7.08 [5.02–9.80]
PSA density [ng/ml/cc] Median [IQR]	0.16 (0.19)	0.10 (0.06)	0.13 (0.07)	0.26 (0.31)
Prostate volume [cc] Median [IQR]	52.2 [36.0–75.3]	59.5 [45.4–94.9]	45.4 [28.9–59.7]	42.4 [33.5–56.9]

IQR, interquartile range; SD, standard deviation.

an NPV of 83.3% for any cancer and 86.7% for csPCa Gleason 3 + 4,no lesions of Gleason ≥ 4+3 were identified.

Overall diagnostic accuracy

Overall, the Likert system had a moderate but non-significant higher specificity than PI-RADS (0.77 *vs* 0.66; χ^2 (1)=3.1; *p* = 0.078) but with unchanged sensitivity (both 0.94) leading to a significantly higher AUC (0.92 *vs* 0.87, *p* = 0.002) and higher PPV (0.66 *vs* 0.58); Table 2, Figure 1. At a positive MRI threshold of score ≥3, 107 lesions were identified on one or both of the classification systems, with 92 by Likert and 106 lesions by PI-RADS scoring. 91 lesions at this cut-off were identified by both classifications whilst 16 patients (8%) were reported differently, with significantly more Likert negative PI-RADS positive (*n* = 15) than Likert positive PI-RADS negative cases (*n* = 1), *p* = 0.001.15 MRIs reported negative by Likert but positive on PI-RADS (11 category 3, three score 4, and one PI-RADS 5), of these six patients underwent biopsy, with five benign and one Gleason 3 + 3 tumour (Supplementary Figure 1).Conversely, one Likert positive (score 4) but PI-RADS negative (score 2) patient had a lesion in the region of the AFMS where targeted biopsy showed Gleason 3 + 3.

Score outcomes at prostate lesion level

Analyzing the individual scores of the two evaluation systems, in 92/199 patients (46.2%) the final score obtained was different; of these 32 relate to PI-RADS score 2 being classified as Likert category 1 all of which showed no cancer. Likert had a higher

proportion of category 5 scores (58/199, 29%) compared to PI-RADS (36/199, 18%). Likert score 5 lesions demonstrated a smaller median diameter compared to PI-RADS 5 at 16 mm and 21 mm, respectively (Supplementary Table 1) but the detection of csPCa was similar at 47/58 (81%) and 29/36 (80.6%), respectively; Table 3, Figure 2. The largest difference was for category 4 lesions, with 53 lesions scored as PI-RADS-4 and only 16 Likert 4 (Figure 3). PI-RADS category 4 lesions were upscored to Likert 5 in 24 cases, and downgraded to either Likert 3 in 15 cases or Likert 2 in 4 cases. With regard to category 4 lesions, the Likert scoring system demonstrated a higher PPV compared to PI-RADS for all cancer (94% *vs* 75%, respectively), and for clinically significant PCa (62% *vs* 53%, respectively). Distribution of Likert and PI-RADS scores by radiologist is provided in Supplementary Table 1.

Score outcomes at prostate zonal level

The majority of the 107 lesions identified were in the peripheral zone (86/107; 80.4%), with 11 in the transition zone (10.3%), 4 in both (3.7%) and 6 lesions arising from the TZ but localized in the AFMS (5.6%); Table 4. Most score changes occurred in PZ lesions (50 target lesions); however, no proportional differences were observed between PZ and TZ (58.1% *vs* 45.5%). In lesions with score ≥3 that differed between Likert and PI-RADS, 12/16 were in the PZ (13.8% overall) and 4/12 in the TZ (23.5% overall), but 3 of the 4 TZ lesions were located in the AFMS; Table 4.
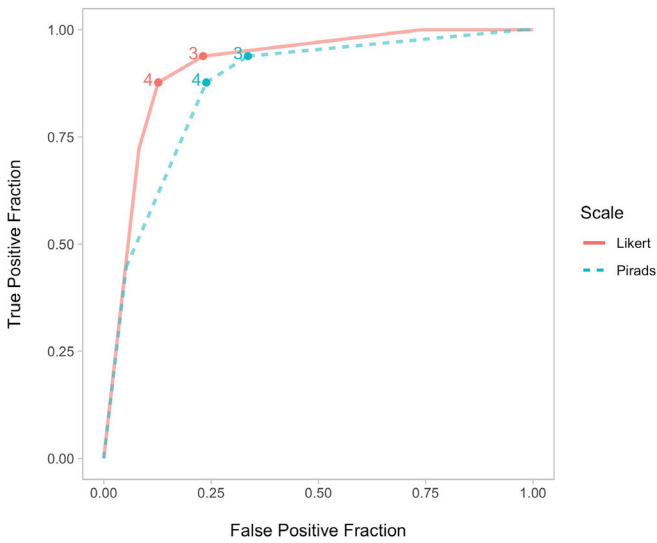
Table 2. Performance comparison of the two systems at a positive MRI cut-off of score ≥3 for csPCa (Gleason score ≥3 +4)

Score	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)	AUC* (CI 95%)
Likert	0.94 (0.85–0.98)	0.77 (0.69–0.84)	0.66 (0.56–0.76)	0.96 (0.91–0.99)	0.82 (0.76–0.87)	0.92 (0.87–0.95)
PI-RADS	0.94 (0.85–0.98)	0.66 (0.58–0.74)	0.58 (0.48–0.67)	0.96 (0.89–0.99)	0.75 (0.69–0.81)	0.87 (0.82–0.92)

AUC, area under the curve; CI, confidence interval; NPV, negative predictive value; PI-RADS, Prostate Imaging–Reporting and Data System; PPV, positive predictive value; csPCa, clinically significant prostate cancer.

p = 0.002 demonstrates significant difference between the two systems.

Figure 1. ROC curves for Likert and PI-RADS scoring. PI-RADS, Prostate Imaging–Reporting and Data System; ROC, receiver operating characteristic.
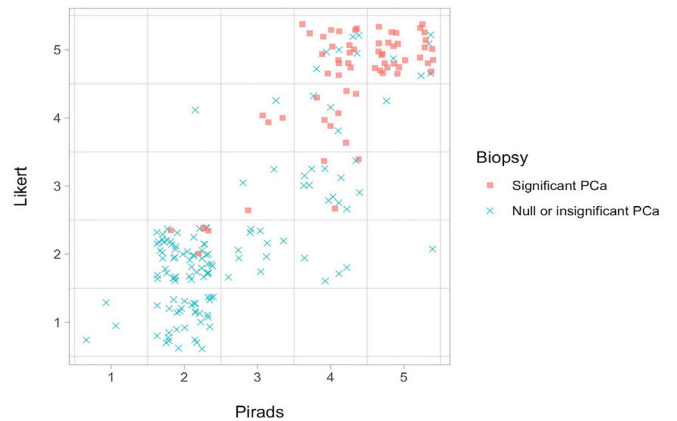

Figure 2. Graphical visualization of outcome prediction change by each scoring system. PCa, prostate cancer.

difference between the scoring systems, and each system only missed one clinically significant cancer. However, Likert scoring had a moderate but non-significant higher specificity based on a lower FP call rate, which could potentially help to reduce the number of unnecessary biopsies performed.

Likert score demonstrated higher PPVs for lesions with an overall higher PPV of 66% compared to PI-RADS at 58% for csPCa and 88 *vs* 76%, respectively for all PCa. Prediction performances were not significantly different between the PZ and TZ for csPCa (χ^2 (3)=4.78, p = 0.189); however, a higher PPV was observed using the Likert scoring system for lesions extending into the AFMS, with a PPV of 75% for Likert *vs* 60% with PI-RADS for csPCa and 100 *vs* 60%, respectively for all prostate cancer; Supplementary Table 1.

DISCUSSION
Our study demonstrates that both Likert and PI-RADS v. 2 are comparable systems for the evaluation of prostate mpMRI, demonstrating high detection rates of clinically significant cancer. At a positive MRI cut-off of score 3, there was an 8%

This study benefited from 3 T MRI scanning interpreted by experienced readers using PI-RADS v. 2, and using an MRI protocol fully compliant with PI-RADS guidelines. The results reported are consistent with those previously reported using both PI-RADS v. 1[9,16] and v. 2,[11,12] showing similar detection rates for csPCa, but with an improvement in diagnostic accuracy for the Likert-based system. Khoo et al recently reported the Likert detected more csPCa across all definitions of clinically significant prostate cancer.[12] Although we report an identical detection rate between the two systems, it is notable that Likert had a moderate increased specificity and significantly higher ROC area under the curve, secondary to a lower positive MRI call rate. This may relate to incorporation of clinical information (including PSA, PSA density clinical symptoms, DRE and family

Table 3. Biopsy outcomes per score

Score		Biopsy		
		Negative	**Gleason 3 + 3**	**Gleason ≥ 3+4**
		N (%)	N (%)	N (%)
1	PI-RADS	3 (100%)	0 (0%)	0 (0%)
	Likert	34 (97%)	1 (3%)	0 (0%)
2	PI-RADS	84 (93%)	2 (2%)	4 (4%)
	Likert	67 (93%)	1 (1%)	4 (6%)
3	PI-RADS	10 (59%)	3 (18%)	4 (24%)
	Likert	8 (44%)	6 (33%)	4 (22%)
4	PI-RADS	13 (25%)	12 (23%)	28 (53%)
	Likert	1 (6%)	5 (31%)	10 (63%)
5	PI-RADS	2 (6%)	5 (14%)	29 (81%)
	Likert	2 (3%)	9 (16%)	47 (81%)

PI-RADS, Prostate Imaging–Reporting and Data System.

Figure 3. Negative PI-RADS score, positive Likert score. 62 year old male, PSA 5.78 ng ml^{-1}. (a) T_2W image, with minor motion artefact. (b): DCE image at the same level. (c/d): DWI shows 12 mm area of restricted diffusion (arrows) in the left anterior stroma region with minor linear, peripheral high signal on b-2000 DWI (c) and low ADC (d). The lesion was scored PI-RADS category 2 initially based on appearances on T_2WI, the dominant sequence. Likert score 4 based on early enhancement at DCE and given restricted diffusion, particularly the high SI on b-2000 images, features which are not in keeping with fibrous tissue. Targeted transperineal biopsy of the lesion showed Gleason 3 + 3 tumour in 2/4 cores, maximally 8 mm. DCE, dynamic contrast enhanced;DWE, diffusion-weightedimaging; PI-RADS, Prostate Imaging–Reporting and Data System; PSA, prostate-specific antigen; SI, signal intensity; T_2WI, T_2weighted imaging.
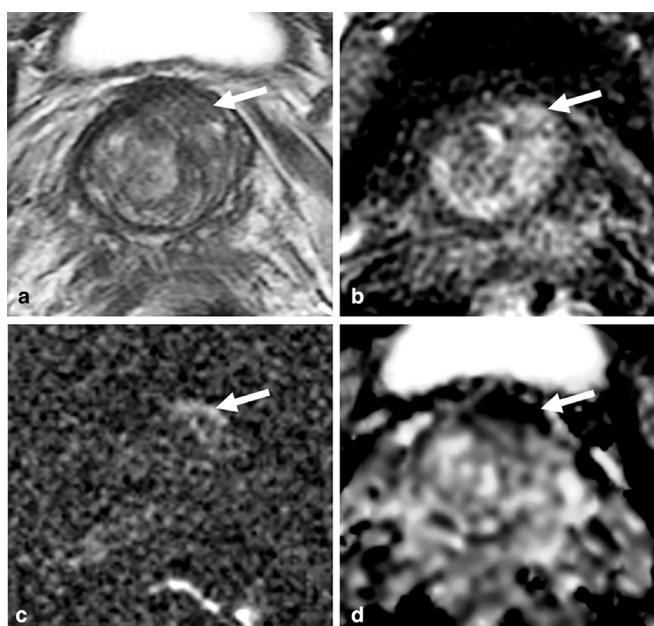
history), particularly for borderline calls at a threshold of Likert category 2–3 and from a clinical standpoint, this resulted in a reduced rate of unnecessary negative biopsies without compromising the detection of csPCa.

The Likert system produced a higher number of category 5 lesions. This is unsurprising given that the PI-RADS system differentiation of category 4 and 5 is based on a size threshold or invasive features. Conversely, the Likert system does not incorporate size criteria and a score 5 is determined based on confidence that a significant lesion is present; our results emphasize this, with Likert-5 lesions being significantly smaller than PI-RADS score 5. However, the percentages of csPCa identified by Likert and PI-RADS category 5 were similar and is supported by previous work suggesting that lesion diameter has only a modest correlation with Gleason grade and significant cancer detection.[17]

Scoring at a lesion level showed the greatest difference for PI-RADS categories 2 and 4. A change from PI-RADS score 2 to Likert-1 will have little if any clinical relevance, and it has been shown that PIRADS v. 2 category 1 prostate classification can be as low as 1%.[18] The lower number of Likert 4 lesions is mainly explained by upscoring of small high probability PI-RADS category 4 lesions to Likert 5, however, there was a notable downscoring of some lesions to category 3. This may relate to PI-RADS score "3 + 1" lesions in the PZ, which can often relate to focal prostatitis,[19] subsequently scored as Likert 2 or 3 with the addition of clinical information. Likert score 4 had a higher PPV compared to PI-RADS 4, and may be explained by the relative heterogeneity of this category incorporating a range from score "3 + 1" through to high probability lesions not quite meeting the category 5 size threshold. The number of Likert/PI-RADS category 3 was similar, with the downgrading of PI-RADS 4 to Likert 3 counterbalanced by some PI-RADS 3 scores being reported as Likert 2.

Rosenkrantz et al[20] found that Likert performed better than PI-RADS v. 1 in the TZ. Assessment of the TZ is more challenging than the PZ, incorporating a subjective assessment of the margins and morphology of the target lesion,[21] thus evaluation of the zone with all sequences in a less rigid manner may be beneficial. Although we did not demonstrate a significant difference in assessment of the TZ, this may relate to limited numbers, and it is worth noting that 75% of lesions within the AFMS region resulted in a change on Likert scoring and with improved outcomes. The updated PI-RADS v. 2.1 may address these issues with a move towards increased use of DWI in the TZ and a separate assessment now proposed for the AFMS region, with emphasis on DCE and high b-value imaging.[22]

Table 4. Prediction change by lesions zone location

	Lesions	Change score	Change results
	N (%)	N (%)	N (%)
PZ	86 (80%)	50 (58%)	12 (14%)
TZ (TZ +AFMS)	17 (16%)	10 (59%)	4 (24%)
TZ (only)	11 (10%)	5 (46%)	1 (9%)
Predominately AFMS	6 (6%)	5 (83%)	3 (50%)
Both (PZ +TZ)	4 (4%)	0 (0%)	0 (0%)
Total	107	60	16

AFMS, anterior fibromuscular stroma.

Our study has some limitations.Likert PI-RADS scores were assigned during the same reading session unblinded, thus the scoring for one system may have biased the scoring of the other. Interobserver agreement could not be assessed due to study design, with each study being assessed by 1 of 4 uro-radiologists as part of a clinical workflow, however, all readers were at the top of their learning curve,[23–25] and all scores ≥3 were reviewed in a multidisciplinary team setting for clinical safety, meaning this may be less relevant.It has been suggested that PI-RADS scoring with its more rigid rules should be used when radiologists are less experienced in mpMRI reporting,[8] thus, the high experience of the readers in this study may mean these results are not fully generalizable. PI-RADS v. 2 was used at the time of the study, however, v. 2.1 is now available and may address some of the issues highlighted here such as assessment of the anterior stroma andinflammatory change in the PZ. The TN rate cannot be fully established in males not undergoing biopsy, and there is potential for sampling error leading to a FN result in cases of both a negative or positive MRI; however, all males underwent a minimum of 14 months of clinical follow-up with at least one repeat PSA. Reassuringly, the NPV of the biopsied patients with a negative MRI was high and no Gleason ≥4 +3 lesions were identified.

In conclusion, both Likert and PI-RADS scoring systems have a high detection of csPCa at a cut-off positive MRI score ≥3. However, Likert scoring by experienced radiologists resulted in lower positive call rate, but with equivalent outcomes, and therefore could help reduce the number of unnecessary biopsies performed.

De long p = 0.002 demonstrates significant difference between the two systems.

REFERENCES

1. Mottet N, van den Bergh RCN, Briers E, et al. EAU – ESTRO – ESUR – SIOG guidelines on prostate cancer 2019. in: European association of urology guidelines.. *European Association of Urology Guidelines Office* 2019;2019 Edition. Vol Edn. presented at the EAU Annual Congress Barcelona 2019. Arnhem, The Netherlands:.

2. Bjurlin MA, Carroll PR, Eggener S, et al. Update of the AUA policy statement on the use of multiparametric magnetic resonance imaging in the diagnosis, staging and management of prostate cancer. *J Urol*Online ahead of print.

3. 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

4. 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

5. 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

6. Brizmohun Appayya M, Adshead J, Ahmed HU, Allen C, Bainbridge A, Barrett T, et al. National implementation of multi-parametric magnetic resonance imaging for prostate cancer detection - recommendations from a UK consensus meeting. *BJU Int* 2018; **122**: 13–25. doi: https://doi.org/10.1111/bju.14361

7. Latifoltojar A, Appayya MB, Barrett T, Punwani S, et al. Similarities and differences between Likert and PIRADS v2.1 scores of prostate multiparametric MRI: a pictorial review of histology-validated cases. *Clin Radiol* 2019; **74**: 895.e1–895.e15. doi: https://doi.org/10.1016/j.crad.2019.08.020

8. National Institute for Health and Care Excellence. Prostate cancer (update) NICE guideline [NG 131]. 2019. Available from: https://www.nice.org.uk/guidance/ng131/documents/draft-guideline [Accessed 16 December 2019].

9. Rosenkrantz AB, Kim S, Lim RP, Hindman N, Deng F-M, Babb JS, et al. Prostate cancer localization using multiparametric MR imaging: comparison of prostate imaging reporting and data system (PI-RADS) and Likert scales. *Radiology* 2013; **269**: 482–92. doi: https://doi.org/10.1148/radiol.13122233

10. Renard-Penna R, Mozer P, Cornud F, Barry-Delongchamps N, Bruguière E, Portalez D, et al. Prostate imaging reporting and data system and Likert scoring system: multiparametric MR imaging validation study to screen patients for initial biopsy. *Radiology* 2015; **275**: 458–68. doi: https://doi.org/10.1148/radiol.14140184

11. Inan I, Aktan A, Ozkanli S, Yildirim A, Aslan A, Gucel S, et al. Comparison of likert and PI-RADS V2 scoring in the diagnosis of prostate cancer. *Annals of Medical Research* 2018; **25**: 651–5. doi: https://doi.org/10.5455/annalsmedres.2018.08.157

12. Khoo CC, Eldred-Evans D, Peters M, Bertoncelli Tanaka M, Noureldin M, Miah S, et al. Likert vs PI-RADS V2: a comparison of two radiological scoring systems for detection of clinically significant prostate cancer. *BJU Int* 2020; **125**: 49–55. doi: https://doi.org/10.1111/bju.14916

13. Barrett T, Turkbey B, Choyke PL. PI-RADS version 2: what you need to know. *Clin Radiol* 2015; **70**: 1165–76. doi: https://doi.org/10.1016/j.crad.2015.06.093

14. Lawrence EM, Tang SYW, Barrett T, Koo B, Goldman DA, Warren AY, et al. Prostate cancer: performance characteristics of combined T_2W and DW-MRI scoring in the setting of template transperineal re-biopsy using MR-TRUS fusion. *Eur Radiol* 2014; **24**: 1497–505. doi: https://doi.org/10.1007/s00330-014-3159-0

15. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, et al. The 2014 International Society of urological pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol* 2016; **40**: 244–52. doi: https://doi.org/10.1097/PAS.0000000000000530

16. Vaché T, Bratan F, Mège-Lechevallier F, Roche S, Rabilloud M, Rouvière O, et al. Characterization of prostate lesions as benign or malignant at multiparametric MR imaging: comparison of three scoring systems in patients treated with radical prostatectomy. *Radiology* 2014; **272**: 446–55. doi: https://doi.org/10.1148/radiol.14131584

17. An JY, Harmon SA, Mehralivand S, Czarniecki M, Smith CP, Peretti JA, et al. Evaluating the size criterion for PI-RADSv2 category 5 upgrade: is 15 mm the best threshold? *Abdom Radiol* 2018; **43**: e44: 3436: . doi: https://doi.org/10.1007/s00261-018-1631-z

18. van der Leest M, Cornel E, Israël B, Hendriks R, Padhani AR, Hoogenboom M, et al. Head-To-Head comparison of transrectal ultrasound-guided prostate biopsy versus multiparametric prostate resonance imaging with subsequent magnetic resonance-guided biopsy in Biopsy-naïve men with elevated prostate-specific antigen: a large prospective multicenter clinical study. *Eur Urol* 2019; **75**: 570–8. doi: https://doi.org/10.1016/j.eururo.2018.11.023

19. Greer MD, Shih JH, Lay N, Barrett T, Kayat Bittencourt L, Borofsky S, et al. Validation of the dominant sequence paradigm and role of dynamic contrast-enhanced imaging in PI-RADS version 2. *Radiology* 2017; **285**: 859–69. doi: https://doi.org/10.1148/radiol.2017161316

20. Rosenkrantz AB, Lim RP, Haghighi M, Somberg MB, Babb JS, Taneja SS, et al. Comparison of interreader reproducibility of the prostate imaging reporting and data system and likert scales for evaluation of multiparametric prostate MRI. *AJR Am J Roentgenol* 2013; **201**: W612–8. doi: https://doi.org/10.2214/AJR.12.10173

21. 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. doi: https://doi.org/10.1016/j.ejrad.2017.08.017

22. Barrett T, Rajesh A, Rosenkrantz AB, Choyke PL, Turkbey B. PI-RADS version 2.1: one small step for prostate MRI. *Clin Radiol* 2019; **74**: 841–52. doi: https://doi.org/10.1016/j.crad.2019.05.019

23. Rosenkrantz AB, Ayoola A, Hoffman D, Khasgiwala A, Prabhu V, Smereka P, et al. The learning curve in prostate MRI interpretation: Self-Directed learning versus continual reader feedback. *AJR Am J Roentgenol* 2017; **208**: W92–100. doi: https://doi.org/10.2214/AJR.16.16876

24. 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. doi: https://doi.org/10.1007/s00330-016-4635-5

25. 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. doi: https://doi.org/10.1111/bju.12892