A semi-automated software program to assess the impact of second reads in prostate MRI for equivocal lesions: results from a UK tertiary referral centre

Alexandre Woernle, Louise Dickinson, Steven Lelie, Doug Pendse, Daniel Heffernan Ho, Navin Ramachandran, Alex Kirkham, Conrad Von Stempel, Shonit Punwani, Chun Wah So, John Withington, Alistair Grey, Thomas Collins, Davide Maffei, Aiman Haider, Alex Freeman, Mark Emberton, Jonathan W Piper, Caroline M Moore, John Hines, Clément Orczyk, Clare Allen, Francesco Giganti

Abstract

Purpose: To investigate the utility of a prostate magnetic resonance imaging (MRI) second read using a semi-automated software program in the one-stop clinic, where patients undergo multiparametric MRI, review and biopsy planning in one visit. We looked at concordance between readers for patients with equivocal scans and the possibility for biopsy deferral in this group.

Methods: We present data from 664 consecutive patients. Scans were reported by seven different expert genitourinary radiologists using dedicated software (MIM®) and a Likert scale. All scans were rescored by another expert genitourinary radiologist using a customised workflow for second reads that includes annotated biopsy contours for accurate visual targeting. The number of scans in which a biopsy could have been deferred using biopsy results and prostate specific antigen density was assessed. Gleason score ≥3 + 4 was considered clinically significant disease. Concordance between first and second reads for equivocal scans (Likert 3) was evaluated.

Results: A total of 209/664 (31%) patients scored Likert 3 on first read, 128 of which (61%) were concordant after second read. 103/209 (49%) of patients with Likert 3 scans were biopsied, with clinically significant disease in 31 (30%) cases. Considering Likert 3 scans that were both downgraded and biopsied using the workflow-second read, 25/103 (24%) biopsies could have been deferred.

Conclusions: Implementing a semi-automated workflow for accurate lesion contouring and targeting biopsies is helpful during the one-stop clinic. We observed a reduction of indeterminate scans after second reading and almost a quarter of biopsies could have been deferred, reducing the potential biopsy-related side effects.

Introduction

Pre-biopsy multiparametric magnetic resonance imaging (MRI) of the prostate increases the detection of clinically significant prostate cancer (csPCa) and minimises that of clinically insignificant disease [1,2]. In the design of both the Prostate Imaging Reporting and Data

Abbreviations: csPCa, Clinically Significant Prostate Cancer; GS, Gleason Score; IQR, Interquartile Range; MCCL, Maximum Cancer Core Length; PI-RADS, Prostate Imaging Reporting and Data System; PSA, Prostate Specific Antigen; PZ, Peripheral Zone; TZ, Transition Zone; DRE, Digital Rectal Examination.

* Corresponding author at: Division of Surgery & Interventional Science, University College London, London, UK

E-mail address: f.giganti@ucl.ac.uk (F. Giganti).
System (PI-RADS) and Likert scoring systems, lesions scoring 3 out 5 (i.e., ‘indeterminate’) are of equivocal likelihood of harbouring csPCa [3] and it has been shown that they tend to harbour csPCa in only a minority of cases [4–6]. However, the decision of whether to biopsy this group of patients remains challenging and other factors must be considered (e.g., prostate specific antigen – PSA – density) [2,7]. PSA surveillance could be a suitable option for avoiding biopsy [8], with evidence supporting that such an approach can be reasonable for up to 89% patients at no curative cost [9]. Efforts have been made to reduce the prevalence of indeterminate lesions in prostate MRI, including the conduct of second reads, where a second radiologist reports the study after assessment by the initial radiologist [10].

Currently, research on this subject has focused on readers with different levels of experience, with inter-reader agreement ranging from 39 to 61% [10–13] and higher positive predictive values for detecting csPCa for more experienced readers [12,13]. However, second reading is a time-consuming task during daily clinical practice given the increasing burden in prostate MRI reporting that radiologists face. Specific tools to facilitate this task and allow the creation of dedicated biopsy plans in a timely and accurate manner would be helpful.

Semi-automated workflows have been tested before to address key issues in prostate MRI, such as image quality, which is a pre-determining factor for the ability to detect csPCa [14], and for patients on active surveillance [15], where such tools optimise the reporting time and allow the creation of clear and standardised structured reports that can be used for targeting biopsies. Equally, semi-automation has been shown to be as accurate in measuring tumour volume as manual methods [16]. At present there are no studies that have investigated the role of a dedicated customised semi-automated workflow to assist second reading in prostate MRI.

The aim of our study was to investigate the added value of a customised semi-automated workflow on the agreement in the number of indeterminate (Likert 3) scans between first and second reads among a group of expert genitourinary radiologists. We also looked at those patients who were biopsied using a structured biopsy plan generated by the workflow and calculated the number of biopsies that could have been subsequently deferred.

2. Materials and methods

2.1. Study population

This is an ongoing prospective study approved and sponsored by the National Health Service North Central London Cancer Alliance. Institutional review board approval at University College London Hospital was obtained and the results included in this study are part of an audit of MR scans and patient data under standard clinical care. All patients who attended our one-stop clinic and underwent a prostate MR scan on the same day between 1st July 2021 and 31st July 2022 were included. Participants who had a prior diagnosis of prostate cancer were excluded.
Fig. 2. Lesion (purple) in the right peripheral zone at midgland contoured by planimetry on axial T2-weighted (A), dynamic contrast enhanced (B), high $b$ value ($b = 2,000$ s/mm$^2$) (C) and apparent diffusion coefficient (D) sequences.

Fig. 3. Second read report and biopsy plan for MR-targeted biopsy generated using the semi-automated workflow.
2.2. One-stop clinic

Patients are referred to the one-stop clinic on an urgent two-week wait pathway by the general practitioner if PSA levels are above the age-specific reference range, or the prostate feels abnormal on digital rectal examination (DRE). Patients receive a multiparametric MRI scan followed by a same-day consultation with a clinical nurse specialist. The results of the scan reported by a dedicated radiologist are explained. Likert scores, clinical data (such as the family history and PSA density), and patient’s preference are factored into a decision of whether to discharge, offer PSA surveillance or offer a biopsy. Where a second opinion for the management plan is necessary (e.g., no visible lesions on MRI but high PSA and PSA density), the case is discussed during our weekly specialist multidisciplinary meeting as per standard of care.

2.3. Image acquisition and analysis

All multiparametric MR scans included in this study were acquired at a field strength of either 1.5 T (Symphony or Avanto, Siemens, Erlangen, Germany) with a high b value of 1,400 s/mm², or 3 T (Achieva or Ingenia, Philips, Best, The Netherlands) with a high b value of 2,000 s/mm². Pelvic surface phased-array coils were used, with no endorectal coil.

MR images were initially reported on the same day by one of seven experienced genitourinary radiologists (LD, DP, DHH, NR, AK, CVS, CA) (all reporting > 1,000 prostate MR scans/year) using a 5-point Likert scale and a semiautomated program for lesion contouring and biopsy planning (MIM® Symphony Dx v. 7.1.2 - Cleveland, OH, USA). In detail, the reporting tool provided a dedicated workflow that allowed the radiologist to look at the different sequences on the same window and contour the prostate (to obtain the volume), seminal vesicles and external urethral sphincter (Fig. 1), and any visible lesion(s) by planimetry (Fig. 2) using a step-by-step procedure. At the end of the workflow, a biopsy plan for MRI-targeted biopsy was created (Fig. 3) and embedded in the final structured report.

Initial readers were aware of presenting PSA and relevant clinical details (e.g., family history of prostate cancer or abnormal DRE) as per standard clinical care of the one-stop clinic. Images were then reviewed within one week by another experienced genitourinary radiologist (FG) (reporting > 1,000 prostate MR scans scans/year) who was not involved in the first read. The second reader was privy to the original report and presenting PSA, but not to the outcome of the consultation with the specialist nurse. This radiologist used a customised version of the semi-automated program for second reads showing the previous contours and had the possibility to upgrade, downgrade, or confirm, as well as to contour any new lesions. A modified biopsy plan that included all the lesions contoured during the first read (even those downgraded on second reads) and any lesion added during the second read was generated and used for biopsy planning.

2.4. Biopsy

Either sedation or local anaesthesia were used on a per patient basis before biopsy, dependent on patient’s choice. Transperineal targeted biopsies were conducted by fellowship-trained surgeons or radiologists.
using a visual-assisted registration no earlier than one week from the scan. Patients whose scans were scored Likert $\leq 2$ but were biopsied because of PSA density $>0.15$ ng/ml received systematic biopsy. In case of negative prostate MR scans (i.e., Likert 1 and 2) and PSA density $\leq 0.15$ ng/ml, patients were discharged to their local general practitioner without biopsy.

2.5. Histopathology

A pool of dedicated consultant histopathologists analysed the biopsy specimens according to the 2019 International Society of Urogenital Pathology guidelines [17], including Gleason score (GS) and maximum cancer core length (MCCL). Rates of csPCa were calculated using a cut-off of $\text{GS} \geq 3+4$ [18].

2.6. Outcomes

We report the results from the use of a customised semi-automated workflow for accurate biopsy planning. For this specific study, when describing a scan using the Likert score, we used only the index lesion (i.e., the lesion with the greatest probability of harbouring csPCa) and excluded secondary lesions. For indeterminate scans, we looked at the number of patients in whom biopsy could have been deferred after second read. This was calculated for patients with downgraded scans to negative (i.e., Likert 1 and 2), no csPCa found on biopsy and a PSA density $\leq 0.15$ ng/ml, as previously described [19–21]. We also determined the concordance between first and second reads in our one-stop clinic, notably for indeterminate scans.

2.7. Statistical analysis

Data are presented as medians and interquartile ranges (IQR) and concordance was calculated as a percentage agreement. Cohen’s kappa statistic ($\kappa$) [22] was calculated for interobserver agreement. Agreement was defined as: slight ($\kappa = 0 – 0.20$), fair ($\kappa = 0.21 – 0.40$), moderate ($\kappa = 0.41 – 0.60$), substantial ($\kappa = 0.61 – 0.80$) and almost perfect ($\kappa = 0.81 – 1.00$). All statistical analysis was performed using IBM SPSS Statistics (version 27.0.1.0, Chicago, Illinois, USA).

3. Results

Over a period of thirteen months, a total of 664 patients underwent prostate MRI during our one-stop clinic. Of these, 580 (87%) were scanned on a 1.5 T system and 84 (13%) on a 3 T scanner. Median age was 65 years [IQR: 59-72]. Median PSA, prostate volume, and PSA density were 5.9 ng/ml [IQR: 4.1-8.9], 52 ml [IQR: 38-77] and 0.11 ng/ml [IQR: 0.07-0.17], respectively.

A total of 285/664 (43%) patients were biopsied, four (1%) of which had no visible lesion on MRI and the decision of systematic sampling was made based on a PSA density $>0.15$ ng/ml. In the biopsied population, 155/285 (54%) harboured csPCa while 16/285 (6%) only clinically insignificant prostate cancer. 114/285 (40%) patients had no cancer on

![Fig. 5. 58-year-old patient with a PSA of 3.2 ng/ml and a prostate volume of 30 cc (PSA density: 0.11 ng/ml). The arrows indicate a midline mid-apical peripheral zone lesion scored Likert 3 both on first and second read (i.e., concordance). This can be seen on T2-weighted (A), dynamic contrast enhanced (B), high b value ($b = 1,400$ s/mm$^2$) (C) and apparent diffusion coefficient (D) sequences. MRI-targeted biopsy revealed Gleason 3 + 4 disease (maximum cancer core length 11 mm).](image-url)
A total of 379/664 (57%) patients did not undergo a biopsy, 10 (3%) of which were scheduled for a biopsy but did not complete it in the study time frame: four patients did not attend or cancelled their biopsy, two were biopsied elsewhere and four were rescheduled due to fitness for biopsy concerns. The remaining 369 patients were either: i) not offered biopsy because of non-visible lesions and a low PSA density (238/369; 64%), ii) opted for PSA surveillance instead (124/369; 34%) or iii) had metastatic disease and were treated directly with hormones and radiotherapy (7/369; 2%).

3.1. Indeterminate scans

A total of 209/664 (31%) patients had equivocal (Likert 3) scans. A total of 103 (49%) patients with Likert 3 findings at first read were biopsied. 66/103 (64%) patients had no cancer on biopsy. PCa was found in 37/103 (36%) cases. 30 (81%) of these tumours were in the peripheral zone (PZ), while 7 (19%) in the transition zone (TZ). CsPCa was found in 31/103 (30%) cases and clinically insignificant prostate cancer in 6/103 (6%).

As shown in Table 1, 78/209 (37%) patients had scans downgraded, 32 of which (41%) were biopsied. This revealed no cancer in 28 cases (88%) but csPCa in 4 cases (12%). Following careful review of the dedicated biopsy plan created using the dedicated semi-automated workflow during the second read, 25/103 (24%) of all patients with equivocal (Likert 3) scans could have had a biopsy deferred (downgrade to Likert ≤ 2/5, no csPCa found on biopsy and a PSA density ≤ 0.15 ng/ml).

3.2. Interobserver agreement

Table 2 displays Likert scores between first and second read. For equivocal (Likert 3) scans, 128/209 (61%) were concordant (κ = 0.60, moderate agreement, 95% CI 0.54-0.67). As far as Likert 4 and 5 scans are concerned, concordance was higher for Likert 4 (87/123; 71%, κ = 0.77, substantial agreement, 95% CI 0.71-0.84) and full (89/89; 100%, κ = 1.00, perfect agreement, 95% CI 1.00-1.00) for Likert 5 scans. In grouping scans as negative (Likert 1–2) and positive (Likert 3–5), concordance was 84% (κ = 0.68, substantial agreement, 95% CI 0.62-0.73). In grouping scans as no suspicious (Likert 1 and 2), equivocal (Likert 3) and suspicious (Likert 4 and 5), concordance was 81% (κ = 0.71, substantial agreement, 95% CI 0.67-0.75). The number of equivocal index lesions decreased from 209 on first read to 152 on second read (27%), as shown in Table 1 and Fig. 4. Examples of concordance and non-concordance are shown in Fig. 5 and Fig. 6.
4. Discussion

In our study, second reads conducted using a semi-automated tool that provides both accurate lesion contouring and a detailed biopsy plan for MR-targeted biopsies were found to decrease the number of indeterminate (Likert 3) scans by almost 30%.

The prevalence of indeterminate (Likert 3) scans in our cohort was 31% but went down to 23% on second read, similar to the reported prevalence of equivocal scans in the literature (22–32%) [6]. Likert is a more subjective scoring system than PI-RADS, the latter not clearly qualifying diffuse changes to the prostate as equivocal. Some studies have described similar rates for csPCa detection between these two scores [23,24], while Khoo et al. [25] found a higher detection rate for Likert. The rate of csPCa in our study was 30%, slightly higher than the values of 21% [6] and 24% [26] that have been reported for PI-RADS scoring. No other study on second reads has calculated the proportion of patients who could have deferred biopsy, hence no comparison is possible.

It is important to note that using the same criteria for biopsy deferral, there would have been four patients in which csPCa would have been missed had the second read been used. Two of these patients (50%) had GS 4 + 3 and GS 3 + 4 with a MCCCL of 3 mm and 5 mm, respectively, but there were severe artefacts on diffusion-weighted imaging, making interpretation extremely difficult. The other two (50%) had GS 3 + 4 disease but a MCCCL < 1 mm. PSA density in all four cases was ≤ 0.15 ng/ml.

For indeterminate scans (Likert 3), concordance was 61%, higher than the 28% found by Ecke et al. [13] and the 26% found by Hansen et al. [11]. This value can be attributed to our study being exclusively performed at a tertiary referral centre and not between referral (i.e., less experienced) and tertiary centres. The same studies [11,13] have also looked at interobserver agreement for describing scans as negative (Likert 1 and 2) and positive (Likert 3 to 5). Concordance was 76% (k = 0.52) [13] and 67% (k = 0.26) [11] were observed, slightly lower than our value of 84% (k = 0.68). For describing scans as not suspicious (Likert 1 and 2), equivocal (Likert 3) or suspicious (Likert 4 and 5), they found values of 58% (k = 0.32) [13] and 46% (k = 0.18) [11], both significantly lower than our value of 81% (k = 0.71).

Our study represents the first in which a dedicated semi-automated workflow has been created and used to streamline the second reading of prostate MR scans in a one-stop clinic and to create a standardised biopsy plan with accurate lesion contouring for MR-targeted biopsies. Our patient group (n = 664) is also the largest at this regard and the fact that this study was observational (downgraded scans still went ahead for biopsy) allowed for our retrospective outcome to be calculated. The results on biopsy deferral might suggest that a reasonable number of patients with indeterminate (Likert 3) scans could defer an invasive procedure that carries the risk of infection, bleeding and overdiagnosis of clinically insignificant prostate cancer.

The limitations of our study must be acknowledged. The first is the design, where there was no blinding between first and second reads. In studies on interobserver agreement, we must consider a tendency for a reviewer to agree with a set of observations [27], especially key here since the readers were both of similar expertise. We did not investigate the value of second reads in the overall cohort (i.e., including also Likert 1, 2 and 4, 5 scores) among readers, as this was not the aim of this study and we believe it would have required a different methodology that takes out the differences between reporters (e.g., in scans reported by a first reader, the second read is done by another reader, and for first reads done by the latter, second reads are done by the previous first reader).

In addition to this, since not all patients were biopsied, we were not able to display the superiority or inferiority of a second read in the ability to rule out csPCa.

Even though the classification used for csPCa is well established, it did not encompass key factors such as the number of positive cores for PCa, the percentage involvement and whether cancer was unilateral or bilateral within the prostate [28].

Lastly, we carried out our analysis using only one dedicated platform, but we acknowledge that there could be other commercially available software with a similar tool.

5. Conclusions

In conclusion, our study suggests that conducting a second read using a customised semi-automated reporting tool that facilitates the creation of a structured report and a biopsy plan for accurate MR-targeted biopsy could potentially defer the number of immediate biopsies for patients with equivocal (Likert 3) scans, allowing patients at low risk to have less frequent procedures, with a positive impact on patient’s quality of life (i.e., reducing the number of side effects).

**Funding:** This study has received funding by the National Health Service North Central London Cancer Alliance.

**Conflict of Interest:** SL is an employee for MIM® Software Inc., Cleveland, OH, USA. JWP is a stockholder and employee for MIM® Software Inc., Cleveland, OH, USA. Both authors did not have any financial or business interests in this study.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

The authors would like to thank Simon Evans, Maria Chu, Jonathan Soriano, Karen Wilkinson, Bethany Penhaligon.

FG is a recipient of the 2020 Young Investigator Award (20YOUN15) funded by the Prostate Cancer Foundation / CRIS Cancer Foundation. FG reports consulting fees from Lucida Medical LTD outside of the submitted work. SP receives research support from the United Kingdom’s National Institute of Health Research (NIHR) UCLH/UCL Biomedical Research Centre. AK is supported by the UCLH/UCL Biomedical Research Centre. ME is a United Kingdom National Institute of Health Research (NIHR) Senior Investigator and receives research support from the UCLH/UCL. NIHR Biomedical Research Centre. CMM is supported by the UKNIHR, Movember, PCUK and the EAU Research Foundation.

**References**


N.L. Hansen, B.C. Koo, F.A. Gallagher, 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 27 (2017) 2259–2266.


L. Marin, M. Ezziane, E. Comperat, et al., Comparison of semi-automated and manual methods to measure the volume of prostate cancer on magnetic resonance imaging, Diagn Interv Imaging 98 (2017) 423–428.


J.F. Zawaldeh, E. Sala, M. Pantelidou, et al., Comparison of Likert and PI-RADS version 2 MRI scoring systems for the detection of clinically significant prostate cancer, Br J Radiol 93 (2020) 20200298.


