Title: The additional value of dynamic contrast enhanced sequences in multiparametric prostate MRI: Data from the PROMIS Study.

Authors

El-Shater Bosaily, A [1,2]

Frangou, E [3]

Ahmed, H. U. [4,5]

Emberton, M [1,6]

Punwani, S [1,6]

Kaplan, R [3]

Brown, LC [3]

Freeman, A [7]

Jameson, C [7]

Hindley. R [8]

Peppercorn, D [8]

Thrower, A [8]

Winkler, M [9]

Barwick, T [9]

Stewart, V [9]

Burns-Cox, N [10]

Burn, P [10]

Ghei, M [11]

Kumaradevan, J [11]

Persad, r [12]

Ash-Miles, J [12]

Shergill, I [13]

Argawal, S [13]

Rosario, D [14]

Salim, F [14]

Bott, S [15]

Evans, H [15]

Henderson, A [16]

Gosh, S [16]

Dudderidge, T [17]

Smart, J [17]

Tung, K [17]

Kirkham, A [6]

On behalf of the PROMIS group.

Corresponding Author: Ahmed El-Shater Bosaily

Clinicalresearchfellow

UCL

Division of surgery and interventional science

University College London

Gower Street

London, WC1E6BT

UNITED KINGDOM

+4407411938521

ashater@nhs.net

Funding

National Institute of Health Research – Health Technology Assessment:

(Project number 09/22/67)

Department of Health Disclaimer:

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the health technology assessment program, NIHR, NHS or the Department of Health.

This project is also supported and partially funded by UCLH/UCL Biomedical Research

Centre and The Royal Marsden and Institute for Cancer Research Biomedical Research Centre and is coordinated by the Medical Research Council Clinical Trials

Unit (MRC CTU) at UCL. It is sponsored by University College London (UCL).

MRC CTU at UCL are supported by an MRC Core grant (MC_UU_12023/28)

The research of Mark Emberton, Hashim Ahmed and Alex Kirkham is supported by core funding from the United Kingdom's National Institute of Health Research (NIHR) UCLH/UCL Biomedical Research Centre.

European Urology abstract

Background:

Multiparametric-MRI (MP-MRI) is established in the diagnosis of prostate cancer, but the need for enhanced sequences has recently been questioned.

Objectives:

To assess whether dynamic contrast enhanced imaging (DCE) improves accuracy over T2 & diffusion sequences.

Design, Setting, and Participants:

PROMIS was a multi-centre, multi-reader trial, with in this part 497 biopsy-naïve men undergoing standardized 1.5T MP-MRIs using T2, diffusion and DCE, followed by a detailed transperineal prostate mapping (TPM) biopsy at 5mmcm intervals. Likert scores of 1-5 for the presence of significant tumour were assigned in strict sequence, for a) T2+ diffusion and then b) T2+ diffusion+DCE images.

Outcome Measurements and Statistical Analysis

For the primary analysis, the primary PROMIS outcome measure (>/=Gleason 4+3 or >/=6mm maximum cancer length) on TPM was used and an MRI score of >/=3 considered positive.

Results and Limitations

Sensitivity without and with DCE was 94% vs 95% , specificity 37% vs 38%, positive predictive value 51% vs 51% and negative predictive value 90% vs 91% respectively (p>0.05 in each case).
The number of patients avoiding biopsy (scoring 1-2) was similar (123/497 vs 121/497, p=0.8). The number of equivocal scores (3/5) was slightly higher without DCE (32% vs 28% p=0.031).

The proportion of MRI equivocal (3/5) and positive (4-5) cases showing significant tumour were similar (23% and 71% vs 20% and 69%).
No cases of dominant Gleason 4 or higher were missed with DCE, compared to a single case with T2+DWI.
No attempt was made to correlate lesion location on MRI and histology, which may be considered a limitation. Radiologists were aware of the patient's PSA.

Conclusions

Contrast adds little when MP-MRI is used to exclude significant prostate cancer.

Patient summary

An iv injection of contrast may not be necessary when MRI is used as a test to rule out significant tumour in the prostate.

Take home message: 30 words. Limit: 40 words:

The addition of DCE did not offer a statistically significant improvement in diagnostic accuracy compared to T2+Diffusion sequences alone, though there was a marginal reduction of equivocal (3/5) MRI results

Article : 2445 words .

Introduction:

Multiparametric Magnetic Resonance Imaging (MP-MRI) is now a well-established tool in the diagnostic pathway of suspected prostate cancer before biopsy[1]. Standards for the conduct of the study have been part of both versions

of the PIRADS scoring systems[2,3], and in each case have included dynamic contrast enhanced (DCE) sequences as standard in addition to T2 and diffusion weighted (DWI) images.

Version 2.1 of PIRADS acknowledges recent data suggesting that the additional utility of DCE for the detection of tumour may be limited, and provides a structure for reporting 'bi-parametric' (T2 and diffusion) studies, though it does not go as far as recommending the routine exclusion of DCE[4]
. One of the main reasons for retaining DCE sequences was the lack of robust data on its diagnostic accuracy from multi-institutional trials with multiple readers[4].

The PROMIS study was a large multi-centre, prospective study with the primary aim of assessing the diagnostic accuracy of pre-biopsy MP-MRI using standard 1.5T machines without an endorectal coil[5]. We report on an embedded prospective assessment of the additional value of DCE over a set of T2 and diffusion-weighted images.

Methods and materials:

The PROMIS trial was a prospective, multi-centre, paired validating cohort study reported to the Standards for Reporting Diagnostic Accuracy (STARD) [5-7]. A total of 576 biopsy naïve men with a clinical suspicion of prostate cancer and PSA <15ng/ml underwent 1.5 Tesla MP-MRI followed by a detailed combined biopsy,

with transperineal mapping of the whole prostate at 5mm intervals as well as the standard 12 core transrectal (TRUS) biopsy. The methods and results are described in detail in a number of papers[5,7,8] but are summarized here.

MRI conduct and reporting:

All patients received an MP-MRI compliant with European Society of Uro-Radiology guidelines[2], with 1.5 Tesla magnetic field strength and a pelvic phased-array coil. This included T2-weighted, diffusion-weighted (including a dedicated b=1400s/mm^2 sequence) and dynamic gadolinium contrast-enhanced (with an approximately 15s time resolution) sequences (Table 1).

Radiologists from the 11 UK centres in the trial all a) had experience of reporting MP-MRI (though there was no minimum criterion for length of experience) and b) attended a single day-long training session. Reporting was on a standardized MRI report format (Figure 1), with the prostate divided into 12 distinct regions of interest. In one sitting, the radiologist assigned a Likert score of 1-5[9] for each grid point, firstly viewing the T2 images alone, then T2+DWI images , and finally T2+DWI+DCE images, in strict sequence and with no retrospective revisions allowed. The overall score of the likelihood of tumour in each patient was defined as the maximum score within the 12 boxes of the grid for each of T2, T2+DWI and T2+DWI+DCE sequences. Reporters were blind to any histological data but were aware of the patient's PSA.

Standard of reference: biopsy

The protocol allowed men with T4 disease on imaging or prostate size >100cc to exit the trial without biopsy; otherwise the series is consecutive. The MP-MRI report remained blinded to all other physicians and trial staff and the combined prostate biopsy procedure was performed with no knowledge of the MP-MRI. The TPM biopsy results were used as the reference standard in the main study findings and also in this paper. Biopsy reporting was completed by one of two expert uropathologists blinded to all MR images and TRUS-biopsy findings. As in the main study two definitions of clinically significant cancerwere used (based on previous work using biopsy simulations[10]): Definition 1 (primary outcome) was Gleason score >/=4+3 or cancer core length >/=6 mm of any grade and Definition 2 was Gleason score >/=3+4 or cancer core length >/=4mm of any grade.

Changes from the pilot phase

For the purpose of this analysis, we included patients from the main phase of the PROMIS trial only (n=497). Seventy-nine patients in the pilot phase were excluded because the sequenced reporting was for a threshold of 'any tumour'. This was amended so that radiologists were asked to determine whether they suspected the presence of 'clinically significant cancer' (>/=0.2cc and/or >/= Gleason 3+4) in the main phase of the trial (see figure 1 for the report form).

Statistical analysis

The diagnostic accuracy of each of the three sequence combinations was assessed against multiple histological thresholds of significant disease using contingency tables. For the primary analysis, a score of 3 or more on MRI was compared to the histological definition 1 of significant tumour , which was that used in the primary outcome paper for PROMIS previously published[5]. PROMIS was not powered to detect differences between sequences, and our analysis must therefore be viewed as exploratory.

Given the paired nature of the data, we used McNemar's test to analyse the differences between T2+DWI and T2+DWI+DCE. Because it was not the main aim of the paper, and to limit the number of statistical comparisons, we did not compare the results of reporting with T2 sequences alone. To compare the positive predictive value (PPV) and negative predictive value (NPV) for the different MRI sequences against TPM-biopsy, we used a general estimating equation (GEE) logistic regression model[11,12],, as these are dependent on prevalence of disease. The TPM results serve as the outcome variable while the explanatory variable is the MRI result for each individual and each sequence. For NPV, the coding logic was reversed (i.e. a negative biopsy was coded as 1 and a positive biopsy was coded as 0) as the test result of interest is correct detection of the absence of clinically significant cancer on the TPM biopsy.

All analyses were done using Stata version 15.1 software (Stata Corporation, College Station, TX, USA)

Results:

In the 497 men assessed, cancer was detected in 71% (354/497) of patients on TPM biopsies. 59% (293/497) had Definition 2 (Gleason score >/=3+4 or cancer core length >/=4mm) and 41% (203/497) had Definition 1 disease (Gleason score >/=4+3 or cancer core length >/=6 mm) (Table 2).

Using Definition 1 (Table 3), the addition of DCE to T2+DWI did not result in statistically significant differences in sensitivity (95% with DCE vs 94% without, p=0.7), specificity (38% vs 37%, p=0.7), PPV (51% vs 51%, p=0.6) and NPV (91% vs 90%, p=0.6), respectively.

When using an alternative threshold of histological significance (definition 2, any tumour of grade ≥3+4 and any tumour of grade ≥4+3) there were no statistically significant differences in diagnostic accuracy metrics between T2+DWI and T2+DWI+DCE (Table 4, p>0.05 in all cases).

The addition of DCE correctly identified all 53 dominant pattern 4 lesions (Table 5) compared to both T2 and T2+DWI which assigned one case as non-suspicious (score 2 compared to score 3 with DCE). Using DCE in combination with T2+DWI, 25% (123/497) of patients were scored negative for significant tumour on MRI, compared to 24% (121/497) with T2+DWI alone (p=0.8, McNemar's test). The addition of DCE slightly reduced the number of equivocal scores (3/5) with 28% of patients classified as equivocal compared to 32% using T2+DWI

alone (p=0.031, McNemar's test) (Table 3). Figure 2 outlines the changes to scores of 3/5 on T2+DWI that were made with the addition of contrast, together with the corresponding histological results.

Figures 3 and 4 show the proportion of significant tumours for each MRI score, using T2+DWI and T2+DWI+DCE, for 4 definitions of clinically significant cancer.

Discussion:

Main findings:

The addition of DCE to T2 and DWI did not improve diagnostic accuracy in a multi-centre study that compared MRI to transperineal mapping biopsy as a reference standard for the detection of clinically significant prostate cancer. The proportion of patients with an equivocal score of 3 was slightly lower for DCE versus non-DCE reporting.

Previous publications.

Three groups in particular have examined a strategy of limited, 'bi-parametric' MRI (T2 and DWI sequences) as a subset of a full PIRADS 2 compliant MP-MRI[13-15]. The methods vary widely, with a DCE time resolution between 3s[14] and 8s[13] and template saturation biopsy [13], targeted biopsy[14] or TRUS biopsy and prostatectomy[15] for histological confirmation, as well as differences in reader experience and reporting criteria. None, however, found a significant

improvement in diagnostic accuracy with contrast. The heterogeneity in methods suggests that a meta analysis will be challenging, although it has been attempted: Woo et al found 20 studies suitable for inclusion in a meta analysis of 'head to head' comparisons, with a total of 2142 patients. They found no convincing difference between bi-parametric and MP-MRI with contrast, although they acknowledged (and attempted to analyse) differences in MRI field strength, use of endorectal coil, reader experience, reporting system (Likert/PIRADS 1/PIRADS2), use of DCE parametric analysis and DCE time resolution [16].

We found that sensitivity of T2 sequences alone for significant tumour was high, and it has been known for some time that most significant prostate cancers are visible on T2 sequences, both in the peripheral[17] and transition zones[18] , but in both previous studies and the current one it was at the expense of a low specificity and a high proportion of equivocal (3/5) scores. Because few centres perform T2 imaging alone, we did not include these results in the statistical comparisons.

Methodological limitations

A number of aspects of the PROMIS study make it particularly relevant to a group of men undergoing MRI as a triage test before biopsy[19]. First, all men with a suspicion of tumour and PSA <15 were included, with only a small number of exclusions due to difficulty performing the biopsy or T4 status.

In addition, the reference standard of transperineal mapping biopsy was applied to all patients without knowledge of the MRI result. This prevents the biases inherent in using MRI-targeted biopsy for confirmation, particularly for a study assessing validity. The study was multi-centre and used local radiologists of varying experience for the primary analysis. The protocol was widely applicable, using 1.5T magnets, no endorectal coil and a feasible DCE time resolution of 15s.

However, there are also potential limitations. First, the analysis was at the level of the prostate: no attempt was made to correlate the position of the tumour on the MRI and TPM biopsy. This has important implications when making inferences about biopsy strategy, but is not relevant when assessing our performance in identifying men at low risk of significant tumour (in other words, in identifying prostates negative for significant cancer). Second, we used a Likert reporting system, which could limit the direct applicability of the study to PIRADS 2 based reporting. However, it may also have helped to detect any potential advantages of DCE, because it allowed the enhanced images (including any morphological criteria that the radiologist deemed useful) to influence the overall suspicion of tumour, rather than just distinguishing between PIRADS 3 and 4 lesions as in the current PIRADS 2 reporting framework[3]. A Likert reporting system has also been recommended for use in the UK by consensus panels[20] and recent National Institute for Health and Care Excellence guidance in the UK[1]. Third, the DCE sequences were biased towards anatomical resolution rather than time resolution, improving image quality but potentially excluding calculation of parameters such as K-trans. However, recent PIRADS 2.1

guidelines acknowledge the lack of data to support a high temporal resolution[4], and PIRADS 2 does not include a recommendation for routine parametric analysis because of a continuing lack of data showing its benefits over interpreting the early enhanced images[3]. Fourth, the PROMIS study was conducted using 1.5T machines, and it is possible that the benefits of dynamic sequences are accentuated with a higher magnetic field strength or the use of an endorectal coil, although two recent studies performed at 3T suggest otherwise[13,14]. Fourth, the radiologist was allowed to know the PSA during reporting. Especially with the potentially subjective analysis of Likert scoring, this information (rather than MRI criteria alone) may have influenced the overall score, in particular inclining the reporter away from a potential 'miss' in the case of a high PSA density. Thus, while knowing the PSA reflects real world practice, it may result in an overestimate of the performance of MRI, whatever the protocol. Finally, the reporting was not supervised, so we cannot absolutely rule out some reporters not looking at and reporting the sequences in the correct order. The importance of doing so was, however, emphasized in written and oral training material. The method of reporting in one sitting enabled a truly prospective study but we cannot eliminate a possible bias from the reporters knowing when they reported the T2 and diffusion sequence that there was a 'final' score using contrast that would be used for the main outcome of the study.

Application to clinical practice.

While the PROMIS study demonstrates the diagnostic accuracy of bi-parametric MRI as a triage test for safely avoiding biopsy, it does not address some important potential benefits of using contrast. There is some evidence that DCE

images improve the measurement of tumour volume[21], although this result is not replicated by all [22]. If true, this may well be because the margins of some tumours are well delineated with contrast and it is possible that it may improve the conduct of targeted biopsies, though no study has attempted to quantify the effect. In addition, there is some evidence that DCE sequences may improve planning for focal therapy[22] or staging tumour at the capsule[23-25]. Finally, there is a consensus that DCE is useful after radiotherapy or ablation[26,27], though the size of the effect when high quality DWI is used has been questioned[28,29]. These potential benefits of contrast must be weighed against its expense and potential risk[30].

Most of the studies on the value of contrast have used experienced readers, but there is some evidence that it helps those with less experience: In a study of 68 selected patients, DCE significantly increased the performance of radiologists with either 100 or 300 cases of reporting experience, but not in those with 1000 cases[31]. This patient group was biased, but even if the result holds in subsequent studies, it is a matter for debate whether a sensible approach is the routine use of contrast, or improved mentorship, specialization and second reads in difficult cases. Similarly, the PIRADS 2.1 document discusses the utility of contrast as a 'safety net' in difficult cases where the other sequences are of sub-optimal quality; it is a matter of debate whether this is routinely included in a scan protocol or used (perhaps in a structure that allows recalls for sub-optimal scans) where it is needed.

Finally, the finding that a lower proportion of patients were given equivocal scores of 3 when contrast was used was statistically significant. However, any clinical impact depends on the way equivocal cases are managed – in particular, whether they undergo biopsy or close surveillance[18].

Conclusion:

The addition of dynamic contrast enhancement did not significantly improve the diagnostic accuracy of T2 + diffusion MRI in a multi-centre, multi-reader study using 1.5T scans. The findings are consistent with recent data from other groups, and allow us to question the necessity of the routine use of contrast in a pre biopsy triage setting, if a high quality MRI is reported by experienced readers.

Acknowledgements: We acknowledge funding from the National Institute of Health Research (NIHR) Health Technology Assessment and Prostate Cancer UK. Mark Emberton's and Alex Kirkham's research is supported by core funding from the United Kingdom's National Institute of Health Research (NIHR) UCLH/UCL Biomedical Research Centre. Hashim Ahmed receives funding fromthe Medical Research Council (UK).

References:

1. NICE Guidance - Prostate cancer: diagnosis and management. *BJU Int* **124,** 9–26 (2019).
2. Barentsz, J. O. *et al.* ESUR prostate MR guidelines 2012. *Eur Radiol* **22,** 746–757 (2012).
3. Weinreb, J. C. *et al.* PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* **69,** 16–40 (2016).
4. Turkbey, B. *et al.* Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* **76,** 340–351 (2019).
5. Ahmed, H. U. *et al.* Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet* **389,** 815–822 (2017).
6. Bossuyt, P. M. *et al.* STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Radiology* **277,** 826–832 (2015).
7. El-Shater Bosaily, A. *et al.* PROMIS — Prostate MR imaging study: A paired validating cohort study evaluating the role of multi-parametric MRI in men with clinical suspicion of prostate cancer. *Contemp Clin Trials* **42,** 26–40 (2015).
8. Brown, L. C. *et al.* Multiparametric MRI to improve detection of prostate cancer compared with transrectal ultrasound-guided prostate biopsy alone: the PROMIS study. *Health Technol Assess* **22,** 1–176 (2018).
9. Likert, R. A technique for the measurement of attitudes. *Arch. Psychology* **22,** 1–55 (1932).
10. Ahmed, H. U. *et al.* Characterizing clinically significant prostate cancer using template prostate mapping biopsy. *J Urol* **186,** 458–464 (2011).
11. Leisenring, W., Alonzo, T. & Pepe, M. S. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics* **56,** 345–351 (2000).
12. Wang, W., Davis, C. S. & Soong, S.-J. Comparison of predictive values of two diagnostic tests from the same sample of subjects using weighted least squares. *Stat Med* **25,** 2215–2229 (2006).
13. Barth, B. K. *et al.* Detection of Clinically Significant Prostate Cancer: Short Dual–Pulse Sequence versus Standard Multiparametric MR Imaging—A Multireader Study. *Radiology* **284,** 725–736 (2017).
14. Kuhl, C. K. *et al.* Abbreviated Biparametric Prostate MR Imaging in Men with Elevated Prostate-specific Antigen. *Radiology* **285,** 493–505 (2017).
15. Junker, D. *et al.* Comparison of multiparametric and biparametric MRI of the prostate: are gadolinium-based contrast agents needed for routine examinations? *World J Urol* **37,** 691–699 (2018).
16. Woo, S. *et al.* Head-to-Head Comparison Between Biparametric and Multiparametric MRI for the Diagnosis of Prostate Cancer: A Systematic Review and Meta-Analysis. *AJR Am J Roentgenol* **211,** W226–W241 (2018).
17. Hricak, H. *et al.* Carcinoma of the prostate gland: MR imaging with pelvic phased-array coils versus integrated endorectal--pelvic

phased-array coils. *Radiology* **193,** 703–709 (1994).
18. Akin, O. *et al.* Transition zone prostate cancers: features, detection, localization, and staging at endorectal MR imaging. *Radiology* **239,** 784–792 (2006).
19. Ahmed, H. U. *et al.* Is it time to consider a role for MRI before prostate biopsy? *Nat Rev Clin Oncol* **6,** 197–206 (2009).
20. Brizmohun Appayya, M. *et al.* National implementation of multi-parametric magnetic resonance imaging for prostate cancer detection - recommendations from a UK consensus meeting. *BJU Int* **122,** 13–25 (2018).
21. Sun, C. *et al.* Comparison of T2-Weighted Imaging, DWI, and Dynamic Contrast-Enhanced MRI for Calculation of Prostate Cancer Index Lesion Volume: Correlation With Whole-Mount Pathology. *AJR Am J Roentgenol* **212,** 351–356 (2019).
22. Cornud, F. *et al.* Tumor Target Volume for Focal Therapy of Prostate Cancer-Does Multiparametric Magnetic Resonance Imaging Allow for a Reliable Estimation? *J Urol* **191,** 1272–1279 (2014).
23. Bloch, B. N. *et al.* Prediction of prostate cancer extracapsular extension with high spatial resolution dynamic contrast-enhanced 3-T MRI. *Eur Radiol* **22,** 2201–2210 (2012).
24. Fütterer, J. J. *et al.* Staging Prostate Cancer with Dynamic Contrast-enhanced Endorectal MR Imaging prior to Radical Prostatectomy: Experienced versus Less Experienced Readers. *Radiology* **237,** 541–549 (2005).
25. Puech, P., Sufana-Iancu, A., Renard, B. & Lemaitre, L. Prostate MRI: Can we do without DCE sequences in 2013? *Diagn Interv Imag* **94,** 1299–1311 (2013).
26. Turkbey, B. *et al.* Correlation of Magnetic Resonance Imaging Tumor Volume with Histopathology. *J Urol* **188,** 1157–1163 (2012).
27. Mason, J., Adiotomre, E., Bownes, P., Carey, B. & Henry, A. Importance of dynamic contrast enhanced magnetic resonance imaging for targeting biopsy and salvage treatments after prostate cancer recurrence. *J Contemp Barchytherapy* **10,** 570–572 (2018).
28. Lotte, R. *et al.* Multiparametric MRI for Suspected Recurrent Prostate Cancer after HIFU:Is DCE still needed? *Eur Radiol* **28,** 3760–3769 (2018).
29. Luzurier, A. *et al.* Dynamic contrast-enhanced imaging in localizing local recurrence of prostate cancer after radiotherapy: Limited added value for readers of varying level of experience. *J. Magn Reson Imaging* **48,** 1012–1023 (2018).
30. Fraum, T. J., Ludwig, D. R., Bashir, M. R. & Fowler, K. J. Gadolinium-based contrast agents: A comprehensive risk assessment. *J Magn Reson Imaging* **46,** 338–353 (2017).
31. Gatti, M. *et al.* Prostate cancer detection with biparametric magnetic resonance imaging (bpMRI) by readers with different experience: performance and comparison with multiparametric (mpMRI). *Abdom Radiol* **44,** 1883–1893 (2019).