Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study

Hashim U Ahmed*, Ahmed El-Shater Bosaily*, Louise C Brown*, Rhian Gabe, Richard Kaplan, Mahesh K Parmar, Yolanda Collaco-Moraes, Katie Ward, Richard G Hindley, Alex Freeman, Alex P Kirkham, Robert Oldroyd, Chris Parker, Mark Emberton, and the PROMIS study group†

Summary

Background Men with high serum prostate specific antigen usually undergo transrectal ultrasound-guided prostate biopsy (TRUS-biopsy). TRUS-biopsy can cause side-effects including bleeding, pain, and infection. Multi-parametric magnetic resonance imaging (MP-MRI) used as a triage test might allow men to avoid unnecessary TRUS-biopsy and improve diagnostic accuracy.

Methods We did this multicentre, paired-cohort, confirmatory study to test diagnostic accuracy of MP-MRI and TRUS-biopsy against a reference test (template prostate mapping biopsy [TPM-biopsy]). Men with prostate-specific antigen concentrations up to 15 ng/mL, with no previous biopsy, underwent 1·5 Tesla MP-MRI followed by both TRUS-biopsy and TPM-biopsy. The conduct and reporting of each test was done blind to other test results. Clinically significant cancer was defined as Gleason score ≥4 + 3 or a maximum cancer core length 6 mm or longer. This study is registered on ClinicalTrials.gov, NCT01292291.

Findings Between May 17, 2012, and November 9, 2015, we enrolled 740 men, 576 of whom underwent 1·5 Tesla MP-MRI followed by both TRUS-biopsy and TPM-biopsy. On TPM-biopsy, 408 (71%) of 576 men had cancer with 230 (40%) of 576 patients clinically significant. For clinically significant cancer, MP-MRI was more sensitive (93%, 95% CI 88–96%) than TRUS-biopsy (48%, 42–55%; p<0·0001) and less specific (41%, 36–46% for MP-MRI vs 96%, 94–98% for TRUS-biopsy; p<0·0001). 44 (5·9%) of 740 patients reported serious adverse events, including 8 cases of sepsis.

Interpretation Using MP-MRI to triage men might allow 27% of patients avoid a primary biopsy and diagnosis of 5% fewer clinically insignificant cancers. If subsequent TRUS-biopsies were directed by MP-MRI findings, up to 18% more cases of clinically significant cancer might be detected compared with the standard pathway of TRUS-biopsy for all. MP-MRI, used as a triage test before first prostate biopsy, could reduce unnecessary biopsies by a quarter. MP-MRI can also reduce over-diagnosis of clinically insignificant prostate cancer and improve detection of clinically significant cancer.

Funding PROMIS is funded by the UK Government Department of Health, National Institute of Health Research–Health Technology Assessment Programme, (Project number 09/22/67). This project is also supported and partly 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).

Introduction

The diagnosis of prostate cancer differs from that in other solid organ cancers where imaging is used to identify those patients who require a biopsy. The prostate cancer diagnostic pathway offers transrectal ultrasound-guided biopsy (TRUS-biopsy) in men who present with an elevated serum prostate specific antigen (PSA). As a result, many men without cancer undergo unnecessary biopsies, clinically insignificant cancers are often detected and clinically significant cancers are sometimes missed.[1,2] TRUS-biopsy also carries significant morbidity and can cause life-threatening sepsis.[3]

A pathway with imaging as a triage test to decide which men with an elevated PSA go on to biopsy might both reduce unnecessary biopsy and improve diagnostic accuracy. Multi-Parametric Magnetic Resonance Imaging (MP-MRI) provides information on not just tissue anatomy but also tissue characteristics such as prostate volume, cellularity, and vascularity. There is some evidence that MP-MRI tends to detect higher risk disease and systematically overlooks low-risk disease,[4,5] which makes it attractive as a potential triage test.[6,7]

In our study, we aimed to investigate whether MP-MRI could discriminate between men with and without clinically significant prostate cancer based on template prostate mapping biopsy (TPM-biopsy) as a reference test. TPM-biopsy is able to accurately characterise disease status in men at risk by sampling the entire prostate

Division of Surgery and Interventional Science, Faculty of Medical Sciences, University College London, London, UK (H Ahmed FRCS, A El-Shater Bosaily MBBCh, Prof M Emberton FRCS); Department of Urology, UCLH NHS Foundation Trust, London, UK (H U Ahmed, A El-Shater Bosaily, Prof M Emberton); Department of Academic Urology, Royal Marsden Hospital, Sutton, UK (C Parker FRCR); MRC Clinical Trials Unit at UCL, London, UK (L C Brown PhD, Prof R Kaplan FRCP, Prof M K B Parmar DPhil, Y Collaco-Moraes PhD, K Ward BSc); Hull York Medical School and Department of Health Sciences, University of York (R Gabe PhD); Department of Urology, Hampshire Hospitals NHS Foundation Trust, UK (R G Hindley FRCS); Department of Histopathology, UCLH NHS Foundation Trust, London, UK (A Freeman FRCPath); Department of Radiology, UCLH NHS Foundation Trust, London, UK (A P Kirkham FRCR); Public and patient representative, Nottingham, UK (R Oldroyd MA)

Correspondence to:
Hashim U Ahmed, Division of Surgery and Interventional Science, University College London, Medical School, Rockefeller Building, 21 University Street, London, WC1E 6DE, UK
hashim.ahmed@ucl.ac.uk

See Online for appendix

Research in context

Evidence before this study

Men with an elevated serum PSA blood test usually undergo transrectal ultrasound guided (TRUS) biopsy. TRUS-biopsy is blind to the location of cancer in the prostate, leading to many men without clinically important cancers undergoing unnecessary biopsy, over diagnosis of clinically unimportant disease, and under-diagnosis of clinically important cancers. In PROTECT, men diagnosed with prostate cancer as a result of PSA screening were randomly assigned to active monitoring, radical prostatectomy, or radical radiotherapy. Results of the study showed no cancer-specific survival differences at a median of 10 years follow-up but reduced time to metastases with treatment. Over three-quarters of men in PROTECT had low risk disease, exemplifying the problem of over-diagnosis from a TRUS biopsy in all strategy and the need to avoid biopsy in such men whilst improving detection of cancer that requires treatment. A systematic review found multi parametric magnetic resonance imaging (MP-MRI) had sensitivity 58–96% and negative predictive value of 63–98% with specificity 23–87%.

MP-MRI has not been assessed using an appropriate reference standard. Previous studies have either used TRUS-biopsy or radical prostatectomy specimens as the reference standard. TRUS-biopsy is inaccurate and radical prostatectomy specimens

are highly selected since men must test positive for cancer on TRUS-biopsy and choose to have surgery so these previous estimates of diagnostic accuracy are potentially inaccurate.

Added value of this study

We assessed the diagnostic accuracy of TRUS-biopsy and MP-MRI using transperineal template prostate mapping biopsies (TPM-biopsies) as the reference standard. TPM-biopsies can be applied to men at risk and are highly accurate since the prostate is sampled every 5 mm. PROMIS shows that MP-MRI has significantly better sensitivity and negative predictive value for clinically important prostate cancer compared with TRUS-biopsy. Thus, MP-MRI could be used as a triage test before first biopsy to allow one quarter of men at risk to avoid biopsy. The lower specificity and positive predictive value of MP-MRI means that a biopsy is still required with a suspicious MP-MRI. Men with suspicious MP-MRI areas can have biopsies guided by these findings. Overall, this strategy could improve detection of clinically important prostate cancer and reduce the number of men diagnosed with clinically unimportant disease.

Implications of all the available evidence

MP-MRI should be used as a triage test before prostate biopsy in men who present with an elevated serum PSA.

every 5 mm. We also aimed to compare the accuracy of MP-MRI with that of TRUS-biopsy.[8] We hypothesised that MP-MRI could be used as a triage test to decide which men with an elevated PSA might safely avoid immediate biopsy.[9]

Methods

PROMIS was a prospective, multi-centre, paired-cohort, confirmatory study, which represented level 1b evidence for diagnostic test assessment[10] and reported to the Standards for Reporting Diagnostic Accuracy.[11] Men were eligible if they had a clinical suspicion of prostate cancer with no previous prostate biopsy. The conduct and reporting of each test was done blind to the other test results. The full details of our protocol can be accessed at http://www.ctu.mrc.ac.uk/our_research/research_areas/cancer/studies/promis/.[8] Ethics committee approval was granted by National Research Ethics Service Committee London (reference 11/LO/0185).

Our primary objectives were to establish the proportion of men who could safely avoid biopsy and the proportion of men correctly identified by MP-MRI to have clinically significant prostate cancer. We also carried out a head-to-head comparison of the accuracy of TRUS-biopsy and MP-MRI in terms of sensitivity, specificity, positive predictive value and negative predictive value or clinically significant prostate cancer, using TPM-biopsy as the reference standard. TPM-biopsy was chosen as the reference test because it samples the entire prostate, is

highly accurate with estimated 95% sensitivity for clinically significant prostate cancers due to its 5 mm sampling frame. Third, the test can minimise selection and work-up biases because it can be applied to men at risk who have had no previous biopsy (appendix figure S1).

Patients

Men who had never had a prostate biopsy were eligible if there was clinical suspicion they might have prostate cancer and they had been advised to have a prostate biopsy. This included men with an elevated serum PSA (up to 15 ng/mL) within previous 3 months, suspicious digital rectal examination, suspected organ confined stage T2 or lower on rectal examination, or family history. Eligible men were aged at least 18 years, fit for general or spinal anaesthesia, and fit to undergo all protocol procedures including a transrectal ultrasound. Men were required to give written informed consent. Patients were excluded if they were using 5-alpha-reductase inhibitors at time of registration or during the previous 6 months; had previous history of prostate biopsy, prostate surgery, or treatment for prostate cancer (interventions for benign prostatic hyperplasia or bladder outflow obstruction were acceptable); had evidence of a urinary tract infection or history of acute prostatitis within the last 3 months; had any contraindication to MRI (eg, claustrophobia, pacemaker, estimated glomerular filtration rate ≤50); had any other medical condition precluding procedures

described in the protocol; or had previous history of hip replacement surgery, metallic hip replacement, or extensive pelvic orthopaedic metal work.

Procedures

Test 1: MP-MRI (index test)

Patients received a standardised MP-MRI, compliant with European Society of Uro-Radiology guidelines, with 1·5 Tesla magnetic field strength and a pelvic phased-array coil. T1-weighted, T2-weighted, diffusion-weighted and dynamic gadolinium contrast-enhanced imaging sequences were acquired (appendix table S1). The protocol allowed men to be withdrawn after the MP-MRI scan if there was evidence of T4 disease or if the prostate volume was greater than 100 mL as TPM-biopsy could not be applied fully to such large prostates. All MRI scanners used by sites and individual MP-MRI scans underwent quality control checks by an independent commercial imaging Clinical Research Organization appointed through open tender (Ixico Ltd, London, UK). Scans deemed of insufficient quality were repeated before the biopsy.

MP-MRI scans were reported at each centre by dedicated urologic radiologists who had previous experience of reporting prostate MP-MRI. They also underwent centralised training involving an initial whole day course, in which 20–30 cases were reviewed individually, scored, and then reviewed as a group. A further training day occurred after the pilot phase with further 20–30 cases reviewed individually and collectively. Radiologists were provided with clinical details including PSA, digital rectal examination findings, and any other risk factors such as family history. A 5-point Likert radiology reporting scale was used to designate prostates as highly unlikely (1), unlikely (2), equivocal (3), likely (4), and highly likely (5) to harbour clinically significant prostate cancer. An MP-MRI score of 3 or greater designated a suspicious scan for the purpose of our primary outcomes. This scoring system was based on the outputs of a consensus group[12] convened before the publishing of the Prostate Imaging and Data Reporting System (PIRADS) MP-MRI reporting consensus.[13] Subsequent comparisons of the Likert and PIRADS reporting schemes have yielded similar results.[14,15] To assess inter-observer agreement, 132 scans from the lead site were re-reported by a blinded second radiologist based at that site.

Tests 2 and 3: combined biopsy procedure

Once the MP-MRI report had been deposited at the central trial office, a combined prostate biopsy procedure was done under general or spinal anaesthesia. Patients and physicians remained blinded to the MP-MRI images and report. Patients first underwent a TPM-biopsy[16,17] followed by TRUS-biopsy. We combined TPM-biopsy with TRUS-biopsy under the same procedure to reduce patient visits and minimise dropout between tests. Due to ethics committee concerns, TRUS-biopsy was done after the TPM-biopsy to minimise infection risk. The

independent Trial Steering Committee monitored safety of this combined procedure in terms of sepsis and other important side-effects, and no concerns were raised during the trial.

The reference test (TPM-biopsy) was done with core biopsies taken every 5 mm and centrally reported at the lead centre (UCLH) by one of two expert uropathologists blinded to all MR images and TRUS-biopsy findings.

In the standard test (TRUS-biopsy), 10–12 core biopsies were taken as per international standards,[18] with each core identified and processed separately. The TRUS-biopsy samples were reported by expert uropathologists at each site blinded to the all MR images and TRUS-biopsy findings.

Definition of clinically significant prostate cancer

Disease significance was defined by criteria previously developed and validated for use with TPM-biopsy for detection of primary Gleason grade 4 or greater[19] and cancer core length predictive for the presence of lesions 0·5 mL or larger.[20–23] Gleason scoring was based on the most frequent pattern and not the highest grade detected on histological analysis. The primary definition used a histological target condition on TPM-biopsy that incorporated the presence of Gleason ≥4 + 3 or more, or a maximum cancer core length (MCCL) involvement of
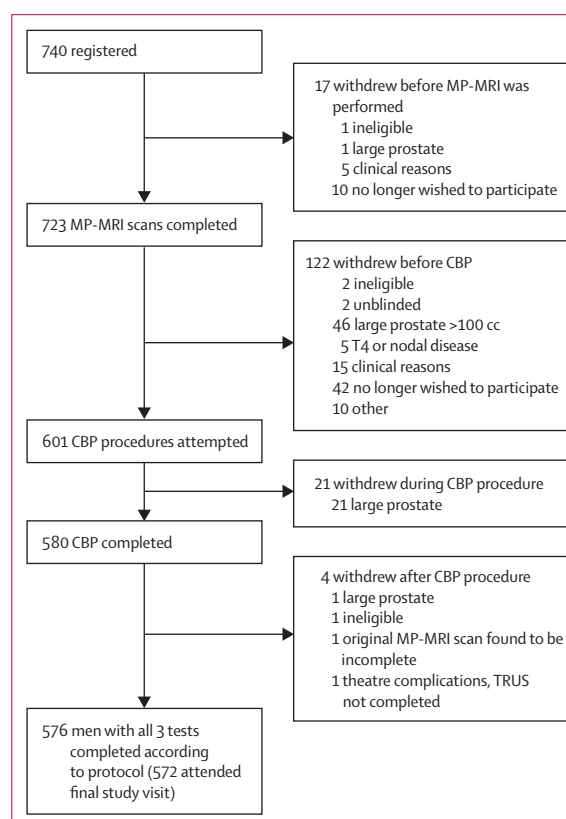

Figure 1: Trial profile
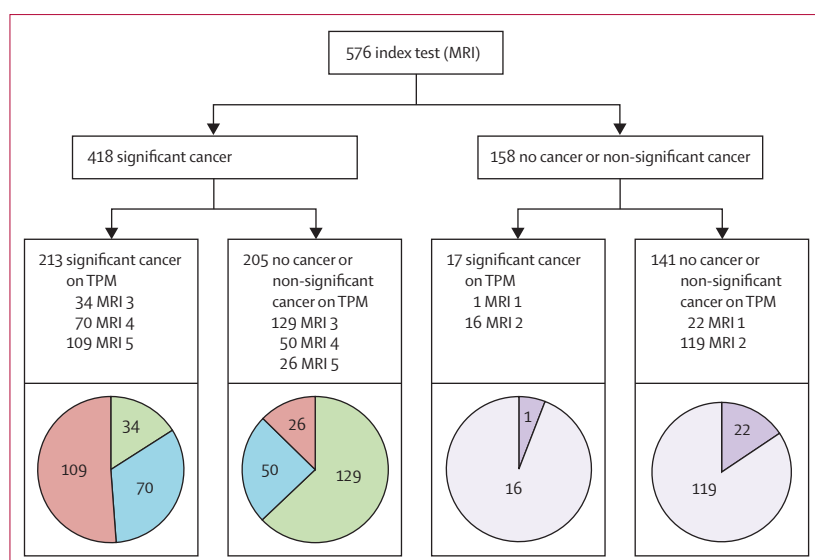MP-MRI=multi-parametric MRI. CBP=combined biopsy procedure.

Figure 2: Diagnostic accuracy for detection of clinically significant cancer (primary definition) between MP-MRI and TPM-biopsy

MP-MRI=multi-parametric MRI. TPM-biopsy=template prostate mapping biopsy. Pie charts represent actual MP-MRI scores 1–5. Sensitivity 93% (95% CI 88–96), positive predictive value 51% (46–56), specificity 41% (36–46), negative predictive value 89% (83–94).
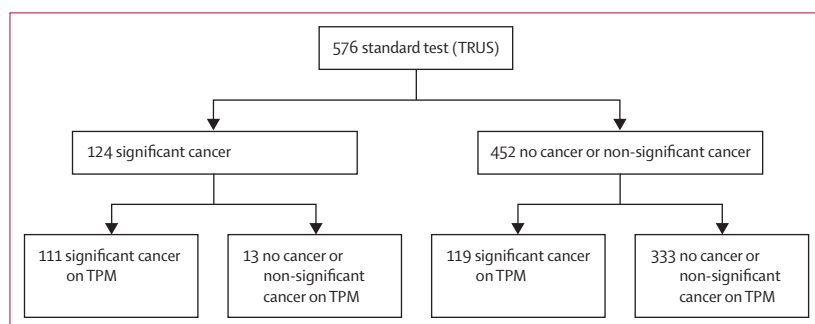

Figure 3: Diagnostic accuracy for detection of clinically significant cancer (primary definition) between TRUS-biopsy and TPM-biopsy

TRUS-biopsy=transrectal ultrasound-guided prostate biopsy. TPM-biopsy=template prostate mapping biopsy. Sensitivity 48% (95% CI 42–55), positive predictive value 90% (83–94), specificity 96% (94–98), negative predictive value 74% (69–78)

6 mm or more in any location. Other definitions of clinical significance were also assessed secondarily.

Sample size

Power calculations were done in relation to precision around the estimates for MP-MRI accuracy in terms of the joint primary outcomes of sensitivity and specificity, a head-to-head comparison of MP-MRI versus TRUS-biopsy, and an assumed underlying prevalence of primary definition clinically significant cancer of 15%. All calculations were based on 90% power and 5% significance (2-sided). This generated a minimum target of 321 (for strong correlation between the tests) and maximum 714 men (for no correlation between the tests).

The Independent Trial Steering Committee carried out an a-priori interim review after 50 men had undergone all 3 tests, and although a higher than anticipated prevalence of any cancer was observed at that time, no changes were recommended to the target sample size.

Statistical analysis

Our sample size target was 714 men. All statistical analyses were done according to a statistical analysis plan agreed before inspection of the data. All analyses were done using Stata version 13.0 software (Stata Corporation, College Station, TX, USA). For each comparison, 2×2 contingency tables were used to present the results and calculate the diagnostic accuracy estimates with 95% confidence intervals. The unit of assessment for our 2×2 contingency table for assessment of accuracy was one patient (ie, the whole prostate). The statistical analysis plan pre-specified that TPM-biopsy results would take precedence over TRUS-biopsy results even if TRUS-biopsy detected clinically significant cancers that TPM-biopsy missed.

Given the paired nature of the test results, McNemar tests were used for the head-to-head comparisons of sensitivity and specificity between MP-MRI and TRUS-biopsy. Given that the positive and negative predictive values are dependent on prevalence of disease, a general estimating equation (GEE) logistic regression model was used to compare the positive predictive value and negative predictive value for MP-MRI and TRUS-biopsy against TPM-biopsy.[24,25] Odds ratios represent the odds of each test correctly detecting the presence or absence of disease. For specificity and negative predictive value, the coding logic is reversed as the correct test result is a negative test result. Ratios are presented as TRUS relative to MP-MRI so ratios greater than 1 favour TRUS and ratios less than 1 favour MP-MRI. LCB had full access to the data and HUA had responsibility for submission of the manuscript.

Results

Between May 17, 2012, and Nov 9, 2015, 740 men were recruited and registered across 11 centres (figure 1). A total of 576 men underwent all 3 tests with 164 withdrawn for various reasons (appendix table S2). Baseline characteristics for all men (both included and withdrawn) are presented in appendix table S3. The median time between MP-MRI and combined biopsy was 38 days (IQR 1–111) days.

Cancer was detected on TPM-biopsy in 408 (71%) of 576 men (95% CI 67–75%). The prevalence of clinically significant cancer according to the primary definition was 230 (40%) of 576 men (36–44). Gleason score ≥4 + 3 occurred in 56 (10%) of 576 men (7–12) and Gleason ≥3 + 4 or greater 220 (38%) of 576 (34–42). 174 men were classified as significant on the basis of core length despite having low grade disease (appendix table S4).

Data collection was more than 95% complete. For 13 men, clinically significant cancer was detected on TRUS-biopsy but missed on TPM. The statistical analysis plan specified that the TPM-biopsy results should take precedence so in 13 (2%) of 576 men in whom TRUS-biopsy designated a patient as having clinically significant cancer, these were treated as false positives as TPM-biopsy found no cancer or clinically insignificant cancer.

Figures 2 and 3 present the diagnostic results for sensitivity, specificity, positive predictive value and negative predictive value for MP-MRI and TRUS-biopsy against TPM-biopsy. Appendix figure S2 presents the proportion of clinically significant disease within each MP-MRI score.

Sensitivity of MP-MRI for clinically significant cancer was 93% (95% CI 88–96%) and negative predictive value 89% (83–94%). Specificity of MP-MRI was 41% (36–46%) with positive predictive value 51% (46–56%). 158 (27%) of 576 men had a negative MP-MRI, of whom 17 had clinically significant cancer on TPM-biopsy (figure 2). All 17 men had Gleason grade 3+4 or less with core lengths that ranged from 6–12 mm (appendix table S5). From figure 3, of the 119 significant cancers missed by TRUS-biopsy, 13 were Gleason 4+3, 99 Gleason 3+4 and 7 Gleason 3+3 (appendix table S5).

MP-MRI was more accurate than TRUS-biopsy in terms of both sensitivity (93% *vs* 48%; McNemar test ratio 0·52 [95% CI 0·45–0·60]) and negative predictive value (89% *vs* 74%, GEE model estimate for odds ratio 0·34 [0·21–0·55]; p<0·0001). TRUS-biopsy showed better specificity (41% *vs* 96%; McNemar test ratio 2·34 [2·08–2·68], p<0·0001) and positive predictive value (51% *vs* 90%; GEE model estimate for odds ratio 8·2 [4·7–14·3], p<0·0001; table).

We considered the implications of using MP-MRI by comparing the standard strategy of TRUS-biopsy for all men to two alternative strategies using MP-MRI as a triage test where only men with a suspicious MP-MRI (Likert score ≥3) would go on to biopsy (appendix table S6). Under the worst case scenario, a standard TRUS-biopsy would be done. Under the best case scenario, the biopsies would be guided by the MP-MRI findings and results are presented assuming targeted biopsies would achieve similar diagnostic accuracy as TPM-biopsy.[26,27] For both these scenarios, 158 (27%) of 576 men would avoid a primary biopsy. For the worst case scenario, an absolute reduction in the over-diagnosis of clinically insignificant cancers might be seen, of 28 (5%) fewer cases per 576 men (relative reduction of 31%, 95% CI 22–42%). For the best case scenario, over-diagnosis of clinically insignificant cancer might be increased to 21%, ie, 31 (5%) more cases per 576 men. For the correct diagnosis of clinically significant cancer, the best case scenario might lead to 102 (18%) more cases of clinically significant cancer being detected per 576 men compared with the standard pathway of TRUS-biopsy for all (table S6). As we did not test MRI targeted TRUS

	MP-MRI, % (95% CI)	TRUS-biopsy, % [95% CI]	Test ratio* [95% CI]	p value
Primary definition (Gleason score ≥4+3 or cancer core length ≥6 mm), prevalence of clinically significant cancer 230 (40%, 36–44%)				
Sensitivity test	93 (88–96)	48 (42–55)	0·52 (0·45–0·60)	p<0·0001
Specificity test	41 (36–46)	96 (94–98)	2·34 (2·08–2·68)	p<0·0001
PPV	51 (46–56)	90 (83–94)	8·2 (4·7–14·3)	p<0·0001
NPV	89 (83–94)	74 (69–78)	0·34 (0·21–0·55)	p<0·0001
Secondary definition (Gleason score ≥3+4 or cancer core length ≥4 mm), prevalence of clinically significant cancer 331 (57%, 53–62%)				
Sensitivity test	87 (83–90)	60 (55–65)	0·69 (0·64–0·76)	p<0·0001
Specificity test	47 (40–53)	98 (96–100)	2·11 (1·85–2·41)	p<0·0001
PPV	69 (64–73)	98 (95–100)	22·7 (8·6–59·9)	p<0·0001
NPV	72 (65–79)	65 (60–70)	0·70 (0·52–0·96)	p=0·025
Any Gleason score 7 (≥3+4), prevalence of clinically significant cancer 308 (53%, 49–58%)				
Sensitivity test	88 (84–91)	48 (43–54)	0·55 (0·49–0·62)	p<0·0001
Specificity test	45 (39–51)	99 (97–100)	2·22 (1·94–2·53)	p<0·0001
PPV	65 (60–69)	99 (95–100)	40·8 (10·2–162·8)	p<0·0001
NPV	76 (69–82)	63 (58–67)	0·53 (0·38–0·73)	p<0·0001

Prevalence of disease on TPM-biopsy, N (%, 95% CI) *McNemar test to compare sensitivity and specificity present ratio of proportions. TPM-biopsy=template prostate mapping biopsy. MP-MRI=multi-parametric-MRI. TRUS-biopsy=transrectal ultrasound-guided prostate biopsy. PPV=positive predictive value. NPV=negative predictive value. General Estimating Equation logistic regression model to compare PPV and NPV present odds ratios. All ratios presented as TRUS relative to MRI.

Table: Diagnostic accuracy of TRUS-biopsy and MP-MRI in the detection of clinically significant prostate cancer using alternative secondary definitions of clinically significant cancer

biopsy, the actual effect of including MP-MRI into the pathway probably lies somewhere between these best and worst case scenarios.

We also evaluated the diagnostic accuracy of TRUS-biopsy and MP-MRI for other definitions of clinical significance on TPM-biopsy. The second definition we used was Gleason ≥3+4 or any grade with cancer core length 4 mm or greater. We also evaluated diagnostic accuracy for the presence of any Gleason score 7 (≥3+4) prostate cancer. The results for all 3 definitions are presented in the table and despite quite different prevalence of disease, the performance of the diagnostic tests did not alter markedly. Clinically significant cases missed under these definitions are presented in appendix tables S5 and S7.

For the 132 men who had blinded, double reporting of their MP-MRI scans, agreement for detection of clinically significant cancer (primary definition) according to the dichotomisation of the MP-MRI scores (1–2 as negative, 3–5 as positive) was 80% (95% CI 72–87%, appendix table S8). This corresponded to a kappa statistic of 0·5 (moderate agreement). PROMIS was done in 11 UK sites with varying experience of MP-MRI reporting. Analysis of the primary outcome results stratified between the central training site (UCLH) and non-UCLH sites reported almost identical findings.

There were 44 reports of Serious Adverse Events during the study (44 [5·9%] of 740 men, 95% CI 4·4–7·9). Of note, 8 (1%) of 576 cases were of sepsis secondary to

urinary tract infection and 58 (10%) of 576 cases of urinary retention. Appendix table S9 gives further details on all side-effects after each test.

Discussion

PROMIS is the first study to our knowledge that presents blinded data on the diagnostic accuracy of both MP-MRI and TRUS-biopsy against an accurate reference test in biopsy-naive men with a suspicion of prostate cancer. It is the largest registered trial to date of the population at risk, across many centres and in which the conduct and reporting of each test was standardised and done blind to the other test results.[28,29] PROMIS represents level 1b evidence for assessment of diagnostic accuracy. The main findings suggest that if MP-MRI was used as a triage test, one-quarter of men might safely avoid prostate biopsy. The high negative predictive value is reassuring in that a negative MP-MRI result implies a high probability of no clinically significant cancer. Further, over-diagnosis of clinically insignificant cancers might be reduced while detection of clinically significant cancers improved compared with the standard of TRUS-biopsy for all men. The lower specificity and positive predictive value of MP-MRI shows that a biopsy, with the needles deployed based on the MP-MRI findings, is still needed in those men with a suspicious MP-MRI.

Our results support the findings of systematic reviews that assess the diagnostic accuracy of MP-MRI.[30,31] The reviews declared sensitivities of 58–96%, negative predictive value of 63–98% and specificity of 23–87%. The ranges were broad because of the single centre nature of the studies, each of which invoked different target conditions on different reference standards. Most studies were limited by retrospective analysis, non-blinding of imaging findings (incorporation and reporting biases), and MP-MRI comparison with inaccurate (TRUS-biopsy) or inappropriate (radical prostatectomy) reference tests. One other prospective study compared MP-MRI with TPM-biopsy that reported interim[32] and then final results.[33] This study reported 96% sensitivity, 36% specificity, 92% negative predictive value and 52% positive predictive value for detection of clinically significant cancer (defined as Gleason score 7–10 with more than 5% Gleason grade 4, 20% or more positive cores, or 7 mm or larger tumour). This Australian study was not blinded, was single-centre, permitted two magnetic field strength scanners (1·5 Tesla or 3·0 Tesla), used a TPM-biopsy protocol that sampled the prostate with fewer cores and did not include the standard test, TRUS-biopsy.[34]

Our study has some limitations. First, although the use of a 5 mm sampling frame of the entire prostate, while too invasive for routine clinical use, offered the precision required for a highly accurate reference test by virtue of its uniform sampling density over the entire prostate gland, this did mean prostates over 100 mL had to be excluded due to template grid size and bony pubic arch interference.[35] Exclusion of large prostates might result in a decrease in the proportion of true negatives within PROMIS. Second, we acknowledge that PROMIS represents a selected group although it is encouraging that men who were subsequently withdrawn from the study did not differ from those who completed the study. Third, the sequence of TPM-biopsy followed by TRUS-biopsy might have contributed to the poor accuracy of the standard test due to swelling, distortion, and tissue disruption. The sequencing was based on patient safety and to preserve the integrity of the reference test. Fourth, by the need for blinding, we did not have targeting of MR-suspicious lesions and cannot accurately assess clinical utility of a MR-targeted biopsy approach. Fifth, although we included some measurement of inter-observer variability, these were between two expert readers. Further work is required to measure the inter-observer variability of expert and non-expert reporters. Last, we acknowledge that likelihood ratios and area under the receiver operating characteristic curves were not part of the pre-specified analysis plan. These metrics provide an overall measure of test performance and clarify the relative strengths and weaknesses of each test, particularly as likelihood ratios are independent of disease prevalence.

The MP-MRI scans in PROMIS were done using 1·5 Tesla magnetic field strength. This was chosen because of its wide availability, and on the assumption that, if benefit was shown, then it was likely to be at least maintained at higher magnetic field strengths where greater signal-to-noise ratios can be achieved.

PROMIS was intentionally designed to be done in multiple sites and it was reassuring that analysis of the results stratified between the central training site (UCLH) and non-UCLH sites reported almost identical findings. Of note in PROMIS, the inter-observer agreement results indicate that there was moderate agreement of MP-MRI scores between two independent radiologists. Our results show that variation usually occurred by 1 point on the Likert scale. Nonetheless, this is an important consideration and highlights the necessity for a robust training programme for radiologists. We have shown that a high quality MP-MRI can be delivered in a multicentre setting in the UK NHS, but all health-care settings will need to consider the capacity issues around access to high-quality MP-MRI, high quality reporting and high quality MRI-targeted biopsies.

Studies using TPM-biopsy as the primary biopsy have previously reported a high prevalence of clinically significant disease.[23,36] This highlights the uncertainty around what constitutes clinical significance.[37] In health-care systems with higher rates of previous PSA testing, the proportion of men with clinically significant prostate cancer, as we have defined it, is likely to be lower. If MP-MRI were used as a triage test in these settings, then the number of men who could potentially avoid a primary biopsy might be higher. Further, the recent ProtecT trial

demonstrates no cancer-specific survival benefit over active monitoring when men are treated with radical prostatectomy or radical radiotherapy, although there was a beneficial effect in reducing time to metastases.[38] With the majority of patients having low risk disease in PROTECT, this adds weight to the need for strategies, such as MP-MRI, that reduce the diagnosis of clinically insignificant cancer and are better able to identify higher risk. In the longer-term, we acknowledge that the volume and type of patients being referred to secondary care could change depending on future validation of potential screening strategies that incorporate risk calculators and biomarkers.[39] Men in PROMIS have consented to follow-up through linkage to central registries for mortality and cancer outcomes so long-term outcomes can be ascertained.

Cost-effectiveness analyses of the PROMIS data are underway and will be reported elsewhere, but the primary outcome data provide a strong argument for recommending MP-MRI to all men with an elevated serum PSA before biopsy. Using MP-MRI as a triage test would reduce the problem of unnecessary biopsies in men who have a low risk of harbouring clinically significant cancer, reduce the diagnosis of clinically insignificant disease and improve the detection of clinically significant cancers.

In conclusion, TRUS-biopsy performs poorly as a diagnostic test for clinically significant prostate cancer. MP-MRI, used as a triage test before first prostate biopsy, could identify a quarter of men who might safely avoid an unnecessary biopsy and might improve the detection of clinically significant cancer.

References
1 Caverly TJ, Hayward RA, Reamer E, et al. Presentation of Benefits and Harms in US Cancer Screening and Prevention Guidelines: Systematic Review. *J Natl Cancer Inst* 2016; **108:** djv436.
2 Abraham NE, Mendhiratta N, Taneja SS. Patterns of repeat prostate biopsy in contemporary clinical practice. *J Urol* 2015; **193:** 1178–84.
3 Loeb S, Vellekoop A, Ahmed HU, et al. Systematic review of complications of prostate biopsy. *Eur Urol* 2013; **64:** 876–92.
4 Turkbey B, Brown AM, Sankineni S, Wood BJ, Pinto PA, Choyke PL. Multiparametric prostate magnetic resonance imaging in the evaluation of prostate cancer. *C A Cancer J Clin* 2016; **66:** 326–36.
5 Mowatt G, Scotland G, Boachie C, et al. The diagnostic accuracy and cost-effectiveness of magnetic resonance spectroscopy and enhanced magnetic resonance imaging techniques in aiding the localisation of prostate abnormalities for biopsy: a systematic review and economic evaluation. *Health Technol Assess* 2013; **17:** vii–xix, 1–281.
6 Valerio M, Willis S, van der Meulen J, Emberton M, Ahmed HU. Methodological considerations in assessing the utility of imaging in early prostate cancer. *Curr Opin Urol* 2015; **25:** 536–42.
7 Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ* 2006; **332:** 1089–92.
8 El-Shater Bosaily A, Parker C, Brown LC, et al. PROMIS Group. 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* 2015; **42:** 26–40.
9 Ahmed HU, Kirkham A, Arya M, et al. Is it time to consider a role for MRI before prostate biopsy? *Nat Rev Clin Oncol* 2009; **6:** 197–206.
10 Phillips B, Ball C, Sackett D, et al. Updated by Howick J. March 2009. http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/ (accessed Nov 29, 2016).
11 Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015; **351:** h5527.
12 Dickinson L, Ahmed HU, Allen C, et al. Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a European consensus meeting. *Eur Urol* 2011; **59:** 477–94.
13 Barentsz J, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol* 2012; **22:** 746–57.
14 Rosenkrantz AB, Lim RP, Haghighi M, Somberg MB, Babb JS, Taneja SS. Comparison of interreader reproducibility of the prostate imaging reporting and data system and likert scales for evaluation of multiparametric prostate MRI. *Am J Roentgenol* 2013; **201:** W612–18.
15 Rastinehad AR, Waingankar N, Turkbey B, et al. Comparison of multiparametric MRI scoring systems and the impact on cancer detection in patients undergoing MR US fusion guided prostate biopsies. *PLoS One* 2015; **10:** e0143404.
16 Hu Y, Ahmed HU, Carter T, et al. A biopsy simulation study to assess the accuracy of several transrectal ultrasonography (TRUS)-biopsy strategies compared with template prostate mapping biopsies in patients who have undergone radical prostatectomy. *BJU Int* 2012; **110:** 812–20.
17 Valerio M, Anele C, Charman SC, et al. Transperineal template prostate-mapping biopsies: an evaluation of different protocols in the detection of clinically significant prostate cancer. *BJU Int* 2016; **118:** 384–90.
18 Heidenreich A, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. *Eur Urol* 2011; **59:** 61–71
19 Stark JR, Perner S, Stampfer MJ, et al. Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3? *J Clin Oncol* 2009; **27:** 3459–64.

20 Kepner G, Kepner J. Transperineal biopsy: analysis of a uniform core sampling pattern that yields data on tumor volume limits in negative biopsies. *Theor Biol Med Model* 2010; **7:** 23.

21 Wolters T, Roobol M, van Leeuwen P, et al. A critical analysis of the tumor volume threshold for clinically insignificant prostate cancer using a data set of a random screening trial. *J Urol* 2011; **185:** 121–25.

22 Stamey T, Freiha F, McNeal J, Redwine E, Whittemore A, Schmid H. Localized prostate cancer. Relationship of tumor volume to clinical significance for treatment of prostate cancer. *Cancer* 1993; **71:** 933–38.

23 Ahmed HU, Hu Y, Carter T, et al. Characterizing clinically significant prostate cancer using template prostate mapping biopsy. *J Urol* 2011; **186:** 458–64.

24 Leisenring W, Alonzo T, Sullivan Pepe M. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics* 2000; **56:** 345–51.

25 Wang W, Davis CS, Soong SJ. Comparison of predictive values of two diagnostic tests from the same sample of subjects using weighted least squares. *Stat Med* 2006; **25:** 2215–29.

26 Radtke JP, Kuru TH, Boxler S, et al. Comparative analysis of transperineal template saturation prostate biopsy versus magnetic resonance imaging targeted biopsy with magnetic resonance imaging-ultrasound fusion guidance. *J Urol* 2015; **193:** 87–94.

27 Kasivisvanathan V, Dufour R, Moore CM, et al. Transperineal magnetic resonance image targeted prostate biopsy versus transperineal template prostate biopsy in the detection of clinically significant prostate cancer. *J Urol* 2013; **189:** 860–66.

28 de Groot JA, Bossuyt PM, Reitsma JB, et al. Verification problems in diagnostic accuracy studies: consequences and solutions. *BMJ* 2011; **343:** d4770.

29 de Rooij M, Hamoen EH, Futterer JJ, Barentsz JO, Rovers MM. Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis. *Am J Roentgenol* 2014; **202:** 343–51.

30 Fütterer JJ, Briganti A, De Visschere P, et al. Can clinically significant prostate cancer be detected with multiparametric magnetic resonance imaging? A systematic review of the literature. *Eur Urol* 2015; **68:** 1045–53.

31 Thompson JE, Moses D, Shnier R, et al. Multiparametric magnetic resonance imaging guided diagnostic biopsy detects significant prostate cancer and could reduce unnecessary biopsies and over detection: a prospective study. *J Urol* 2014; **192:** 67–74.

32 Thompson JE, Moses D, Shnier R, et al. Multiparametric magnetic resonance imaging guided diagnostic biopsy detects significant prostate cancer and could reduce unnecessary biopsies and over detection: a prospective study. *J Urol* 2014; **192:** 67–74.

33 Thompson JE, van Leeuwen PJ, Moses D, et al. The Diagnostic Performance of Multiparametric Magnetic Resonance Imaging to Detect Significant Prostate Cancer. *J Urol* 2016; **195:** 1428–35.

34 El-Shater Bosaily A, Arya M, Punwani S, et al. Re: Multiparametric magnetic resonance imaging guided diagnostic biopsy detects significant prostate cancer and could reduce unnecessary biopsies and over detection: a prospective study. *J Urol* 2014; **192:** 67–74.

35 Crawford ED, Rove KO, Barqawi AB, et al. Clinical-pathologic correlation between transperineal mapping biopsies of the prostate and three-dimensional reconstruction of prostatectomy specimens. *Prostate* 2013; **73:** 778–87.

36 Bittner N, Merrick GS, Bennett A, et al. Diagnostic performance of initial transperineal template-guided mapping biopsy of the prostate gland. *Am J Clin Oncol* 2015; **38:** 300–03.

37 Valerio M, Anele C, Bott SR, et al. The prevalence of clinically significant prostate cancer according to commonly used histological thresholds in men undergoing template prostate mapping biopsies. *J Urol* 2016; **195:** 1403–08.

38 Lane JA, Donovan JL, Davis M, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol* 2014; **15:** 1109–18.

39 Grönberg H, Adolfsson J, Aly M, et al. Prostate cancer screening in men aged 50–69 years (STHLM3): a prospective population-based diagnostic study. *Lancet Oncol* 2015; **16:** 1667–76.