
Maarten de Rooij
Multiparametric MRI in prostate cancer
Diagnostic accuracy and economic evaluation

Multiparametric MRI in prostate cancer

Diagnostic accuracy and economic evaluation

Maarten de Rooij

The research in this thesis was supported by a personal Dutch Cancer Society grant of prof. dr. J.O. Barentsz and a personal Radboud university medical center Hypatia fellowship of prof. dr. M.M. Rovers. The work presented in this thesis was carried out within the Radboud Institute for Health Sciences, Radboud university medical center, Nijmegen, The Netherlands; Wesley Medical Centre, Brisbane, Australia; Achmea Kenniscentrum, The Netherlands.

Publication of this thesis was financially supported by the Radboud university medical center, Nijmegen, The Netherlands.

ISBN
978-94-92303-16-5

Cover/design
Ilse Schrauwers, isontwerp.nl, Eindhoven, The Netherlands

Lay-out and printing
Gildeprint, Enschede, The Netherlands

Multiparametric MRI in prostate cancer

Diagnostic accuracy and economic evaluation

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 6 juli 2017
om 12.30 uur precies

door
Maarten de Rooij
geboren op 10 oktober 1983
te Tilburg

Promotoren

Prof. dr. M.M. Rovers

Prof. dr. J.O. Barentsz

Manuscriptcommissie

Prof. dr. P.F.A. Mulders, voorzitter

Prof. dr. M.A.A.J. van den Bosch (UMC Utrecht)

Prof. dr. M.G.M. Hunink (Erasmus MC, Rotterdam & Harvard University, Boston)

TABLE OF CONTENTS

1

Introduction and outline of thesis

INTRODUCTION

Prostate cancer (PCa) is the most common malignancy among men in developed countries, and has a large societal burden.[1] PCa affects 1 out of 7 men, and results in death in 1 out of 39 men.[2] Approximately 50-70 % of the detected PCa is a low-risk type that will never cause symptoms during a lifetime.[3,4] An often-heard expression about this type of PCa is: "you will die with PCa rather than of PCa." However, in the other 30-50% of the men, an intermediate- or high-risk type is diagnosed, which needs to be treated. In the beginning, these more aggressive tumour types often cause no symptoms either, but can eventually lead to PCa progression, with severe symptoms and subsequent early death. The challenge is to identify the aggressive tumours early and to treat these accordingly, while leaving alone low-risk PCa.

Systematic transrectal ultrasound guided prostate biopsy (TRUS-GB) is the current standard of care after an elevated serum prostate specific antigen (PSA). It comprises an invasive procedure, and it is known that important tumours are being missed or misclassified with this method, whereas insignificant tumours which do not require treatment may be found (see Figure 1). Magnetic resonance imaging (MRI), used as an imaging technique before biopsy, shows promising results in improving PCa diagnosis.[5,6] Because of these promising results, the Dutch Prostate Cancer Foundation hopes that prostate MRI will be included in the standard diagnostic pathway for men with a suspicion on PCa. However, he evidence regarding the effectiveness of prostate MRI on quality of life and its related cost-effectiveness is lacking. Another reason is the current lack of evidence of superiority of an MRI guided diagnostic pathway over a TRUS-GB pathway in terms of patient outcome. Furthermore, MRI might not be available in all hospitals. These factors may be a possible explanation for the fact that many health insurance companies still withhold reimbursement for prostate MRI as the (first) diagnostic step after an elevated PSA, and that an MRI before biopsy is still not recommended in the Dutch PCa guidelines.[7]

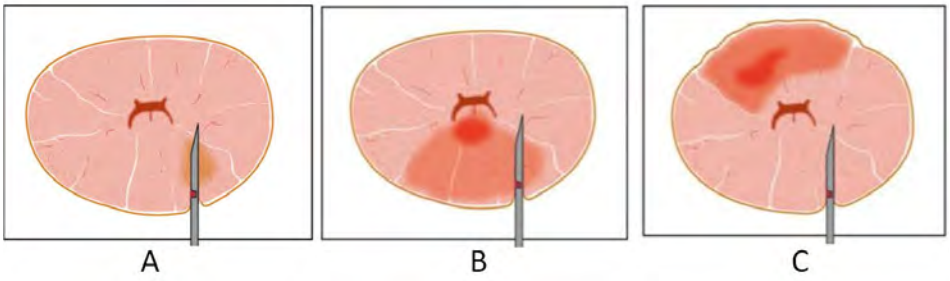


Figure 1 Disadvantages of transrectal ultrasound guided biopsies. Low-risk tumours are detected by chance (A), the most aggressive part of intermediate- and high-risk cancer is not detected (B), and intermediate- and high-risk tumours are missed due to under-sampling of apex, base, and the anterior part of the prostate (C).

CURRENT DIAGNOSTIC TOOLS

As mentioned above, systematic TRUS-GB is the standard diagnostic tool to detect tumours in men with an elevated PSA or abnormal digital rectal examination (DRE).[8] The urologists that perform TRUS-GB take 10 to 12 prostate biopsies under ultrasound guidance in the areas that are most common for PCa. This is the diagnostic standard for many years and gives an excellent guidance of the prostate size and boundaries, but provides limited information on areas suspicious for PCa.

Recent statements from the European Association of Urology (EAU), the U.S. Preventive Screening Task Force (USPSTF), and the American Urologic Association (AUA) critically question the current diagnostic pathway and recommend against the use of widespread PSA testing.[9,10] In this so-called 'PSA dilemma', the harms in the diagnostic pathway outweigh the benefits for most men due to several reasons. First, the PSA test can lead to unnecessary TRUS-GB because 2 out of 3 men with an elevated PSA will have a non-malignant condition like benign prostate enlargement or inflammation of the prostate.[11] Normal PSA levels, on the other hand, do not exclude the presence of PCa.[12] Second, with DRE many tumours are being missed, and the same prostate will often be interpreted differently among different urologists.[13] Figure 1 shows the main disadvantages of TRUS-GB. Because TRUS-GB provides no detailed images of the prostate, low-risk tumours are detected by chance (Figure 1A), and the aggressiveness of intermediate- or high-risk cancers cannot be accurately assessed (Figure 1B).[11,14,15] Furthermore, important tumours are being missed, because harder to reach parts of the prostate are under-sampled, as shown in Figure 1C.[16] The limitations of the current diagnostic pathway lead to inadequate assessment of the tumour type. This

can lead to unnecessary diagnosis of low-risk tumours with subsequent invasive treatment and misdiagnosis and under-treatment of intermediate- to high-risk PCa.[17]

MRI AND MR GUIDED BIOPSY

In the last decades, technology has drastically evolved. New imaging techniques, like advanced ultrasound, ^{11}C-Choline PET/CT, ^{68}Gallium-PSMA PET/CT, whole body MRI, Combidex-MRI, and multiparametric MRI (mp-MRI) have been developed and tested. New tests for molecular biomarkers in urine (PCa antigen 3 (PCA3)) and epigenetic assays (such as ConfirmMDx) also show interesting results.[18,19] To improve the detection of PCa, mp-MRI is currently the best studied imaging technique. Many studies have demonstrated the potential of mp-MRI to solve the aforementioned PSA dilemma. When performed before biopsy, mp-MRI could reduce the detection of insignificant cancer, improve the detection of intermediate- to high-risk PCa without biopsy, and enable targeted biopsy and focal therapy. [6,20,21]

The state-of-the-art prostate mp-MRI combines ordinary anatomic imaging (T2-weighted imaging; T2WI) with more advanced functional imaging techniques.[22] The additional functional techniques show microscopic motion of water (diffusion weighted imaging; DWI), handling of intravenous contrast media (dynamic contrast-enhanced MRI; DCE-MRI), and estimates for metabolism in prostate tissue (proton MR spectroscopic imaging; MRSI). The techniques are used to differentiate normal from malignant prostate tissue and to estimate the aggressiveness of the tumor.[23-25]

The European Society of Urogenital Radiology (ESUR) and the American College of Radiology (ACR) recommend the use of anatomic T2WI combined with two additional functional techniques (DWI and DCE-MRI) for the detection of prostate cancer, with MRSI as optional functional technique.[22,26] The ESUR guideline also describes criteria to report the suspicion of PCa on mp-MRI. These criteria, called Prostate Imaging Reporting and Data System (PI-RADS), are used to generate scores between 1 and 5 for each imaging technique, where 1 = tumour highly unlikely, and 5 = tumour highly likely. The goal of PI-RADS is to provide a standardized 'language' and to allow for risk-stratification between men with benign lesions which do not require biopsy, and men that need further diagnostic workup. In figure 2, an example of mp-MRI of the prostate is shown using T2WI, DWI, and DCE-MRI.

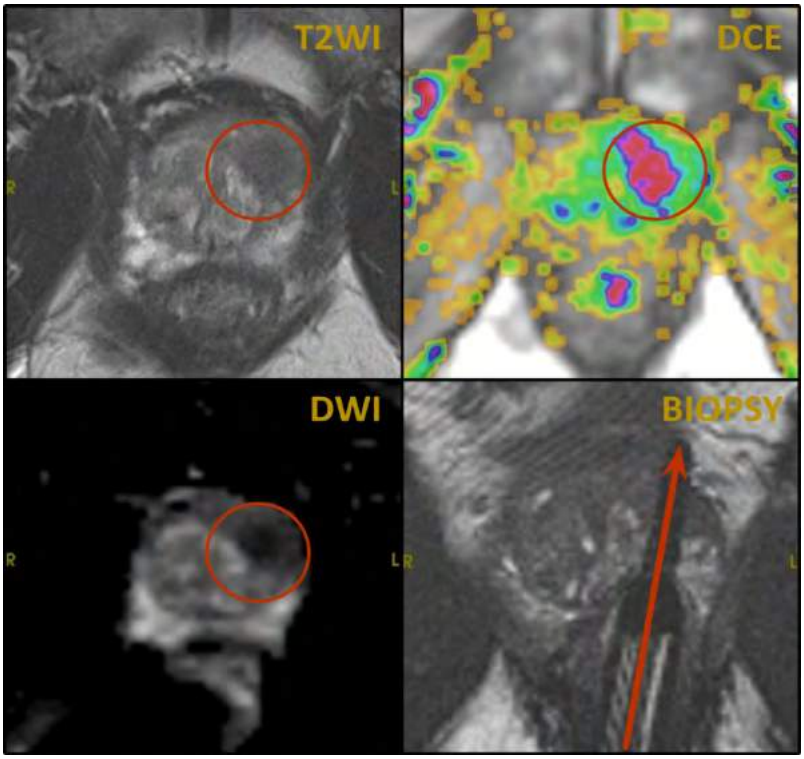


Figure 2 An example of a multiparametric MRI of the prostate of a 66-years old man with an elevated PSA and 2 previous negative TRUS-GB sessions. The images show a suspicious area in the left anterior part of the prostate (red circle). The upper left image shows an anatomical image of the prostate (T2-weighted imaging; T2WI). In the upper right corner, a colour map is shown of the wash-out of contrast medium (dynamic contrast enhanced imaging; DCE-MRI). In the lower left a diffusion weighted image of the prostate is shown (DWI). The black area represents an area with less microscopic motion of water molecules, which is suspicious for malignant tissue. MR guided biopsy of this area (lower right) proved a high risk tumour, which needs active treatment.

Where TRUS-GB is normally performed in a standard set of 10-12 areas of the prostate, mp-MRI enables targeted biopsy in only those areas considered to be suspicious for PCa. Targeted biopsy techniques can, therefore, use fewer biopsies in fewer men to obtain the same rate of intermediate- or high-risk PCa, and can potentially reduce the diagnosis of low-risk PCa.[20] Three types of targeted biopsy techniques are currently being used; (1) cognitive fusion biopsy, (2) MR/US fusion biopsy, and (3) in-bore MR-guided biopsy (MRGB). Cognitive fusion biopsy is a technique in which the urologist performs TRUS-GB in areas with a suspicion of PCa on prior mp-MRI, without visual feedback on the correct biopsy location. MR/US fusion is a biopsy technique that is used by the urologist to fuse the prior images from mp-MRI with real-time ultrasound images to enable targeted biopsy. The in-bore MRGB technique, which is shown in Figure 2, is performed by a radiologist or urologist in the MR device with real-time MR image feedback. This technique is more

time consuming and requires radiologic expertise, but can potentially provide the most accurate results.

EVALUATION OF NEW DIAGNOSTIC TECHNIQUES

For the evaluation of new diagnostic technologies, like mp-MRI and MRGB ('MRI pathway'), it is important to acknowledge the six hierarchical levels of evidence that are stated in the model of Fryback and Thornbury.[27] The first four levels comprise technical performance (1), diagnostic accuracy (2), and impact of the new technology on diagnosis (3) and treatment planning (4). The last two levels focus on patient health outcomes (5) and economic modelling to assess the societal value (6) of the new diagnostic test. Many studies on mp-MRI and MRGB provided evidence on the first four levels. In current literature, however, there is a lack of studies that provide an overview of the evidence on the diagnostic accuracy (level 2). For urologists and radiologists, this overview is essential, because the MRI techniques evolve fast and many different protocols are being used, with a wide variety of reported outcomes. Also, there is a lack of studies that provide evidence on the last two levels, which focus on patient health outcomes (level 5) and on an economic model to assess the societal value of the MRI pathway (level 6). An economic model integrates data from the first 5 levels of evidence and is increasingly important to inform policy makers on new diagnostic techniques, especially in the current trend towards more sustainably and (cost-)effective health care. If new techniques, like the MRI pathway, perform better than the standard of care in clinical outcomes, studies that provide evidence on level 5 and 6 are essential before urologists will use the MRI pathway or other novel techniques as new standard of care.

AIM OF THIS THESIS

The overall aim of this thesis is to evaluate the diagnostic MRI pathway in men with a suspicion of PCa. More specifically this thesis aims:

1. to assess the diagnostic accuracy of mp-MRI for the detection and local staging of PCa.
2. to assess the clinical effectiveness of the MRI pathway in patients with a suspicion of PCa.
3. to assess the cost-effectiveness of the MRI pathway in patients with a suspicion of PCa.

OUTLINE OF THIS THESIS

Part 1 (Chapter 2-4) provides an overview of the current evidence regarding the accuracy of MRI. We studied the role of mp-MRI for the detection of PCa in two separate diagnostic meta-analyses; the first on the accuracy of MRI for the detection of all types of PCa (**Chapter 2**), the second specified on clinically significant PCa and the role of standardized reporting systems (**Chapter 3**). In **Chapter 4,** we systematically reviewed the evidence on the accuracy of MRI for local staging of PCa.

In Part 2 (Chapter 5 and 6), we studied the costs and cost-effectiveness of the current diagnostic pathway and the new MRI pathway. The diagnostic meta-analysis of Chapter 2 has been used as input for a decision analytic model (**Chapter 5***)* that compares the MRI pathway with the current standard of care for men with a suspicion of PCa. In **Chapter 6,** we provide insight in the health care costs of PCa treatment related urinary incontinence costs using a health insurance database.

To further improve the evidence on the role of MRI in the diagnostic pathway, we performed a prospective study, which is described in Part 3 (**Chapter 7**). We studied the clinical outcomes, that is, the detection of low-, intermediate-, and high-risk PCa of both the current standard of care and the new MRI pathway.

In Part 4 (**Chapter 8 and 9**), the main findings of this thesis are discussed and we will give an overview of the most important challenges for the different stakeholders that are involved in the PCa diagnostic pathway. Furthermore, we give recommendations and future directions regarding the possible implementation of the new MRI pathway.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA: a cancer journal for clinicians* 2016; 66(1): 7-30.
2. American Cancer Society Prostate cancer facts. Cancer facts & figures 2009. Atlanta: American Cancer Society; 2009. p. 19-20.
3. Herget KA, Patel DP, Hanson HA, Sweeney C, Lowrance WT. Recent decline in prostate cancer incidence in the United States, by age, stage, and Gleason score. *Cancer medicine* 2016; 5(1): 136-41.
4. The 2016 Cancer System Performance Report. Canadian Partnership Against Cancer. Toronto (ON): The Canadian Partnership Against Cancer; 2016.
5. Futterer 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(6): 1045-53.
6. Schoots IG, Roobol MJ, Nieboer D, Bangma CH, Steyerberg EW, Hunink MG. Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. *Eur Urol* 2015; 68(3): 438-50.
7. Prostaatcarcinoom; landelijke richtlijn versie 2.1. Integraal Kankercentrum Nederland; 2016.
8. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* 2016.
9. Carter HB, Albertsen PC, Barry MJ, et al. Early detection of prostate cancer: AUA Guideline. *The Journal of urology* 2013; 190(2): 419-26.
10. Moyer VA, Force USPST. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine* 2012; 157(2): 120-34.
11. Bonekamp D, Jacobs MA, El-Khouli R, Stoianovici D, Macura KJ. Advancements in MR imaging of the prostate: from diagnosis to interventions. *Radiographics : a review publication of the Radiological Society of North America, Inc* 2011; 31(3): 677-703.
12. Bul M, van Leeuwen PJ, Zhu X, Schroder FH, Roobol MJ. Prostate cancer incidence and disease-specific survival in men participating in the ERSPC with an initial PSA less than 3.0 ng/mL. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011; 29(7_suppl): 7.
13. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *The Cochrane database of systematic reviews* 2013; (1): CD004720.
14. Hambrock T, Hoeks C, Hulsbergen-van de Kaa C, et al. Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort. *Eur Urol* 2012; 61(1): 177-84.
15. Ukimura O, Coleman JA, de la Taille A, et al. Contemporary role of systematic prostate biopsies: indications, techniques, and implications for patient care. *Eur Urol* 2013; 63(2): 214-30.
16. Moosavi B, Flood TA, Al-Dandan O, et al. Multiparametric MRI of the anterior prostate gland: clinical-radiological-histopathological correlation. *Clinical radiology* 2016; 71(5): 405-17.
17. Loeb S, Bjurlin MA, Nicholson J, et al. Overdiagnosis and overtreatment of prostate cancer. *Eur Urol* 2014; 65(6): 1046-55.
18. Van Neste L, Partin AW, Stewart GD, Epstein JI, Harrison DJ, Van Criekinge W. Risk score predicts high-grade prostate cancer in DNA-methylation positive, histopathologically negative biopsies. *The Prostate* 2016; 76(12): 1078-87.
19. Hendriks RJ, van Oort IM, Schalken JA. Blood-based and urinary prostate cancer biomarkers: a review and comparison of novel biomarkers for detection and treatment decisions. *Prostate cancer and prostatic diseases* 2016.
20. Moore CM, Robertson NL, Arsanious N, et al. Image-guided prostate biopsy using magnetic resonance imaging-derived targets: a systematic review. *Eur Urol* 2013; 63(1): 125-40.

21. Siddiqui MM, Rais-Bahrami S, Truong H, et al. Magnetic resonance imaging/ultrasound-fusion biopsy significantly upgrades prostate cancer versus systematic 12-core transrectal ultrasound biopsy. *Eur Urol* 2013; 64(5): 713-9.
22. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* 2016; 69(1): 16-40.
23. Vos EK, Kobus T, Litjens GJ, et al. Multiparametric Magnetic Resonance Imaging for Discriminating Low-Grade From High-Grade Prostate Cancer. *Investigative radiology* 2015; 50(8): 490-7.
24. Vos EK, Litjens GJ, Kobus T, et al. Assessment of prostate cancer aggressiveness using dynamic contrast-enhanced magnetic resonance imaging at 3 T. *Eur Urol* 2013; 64(3): 448-55.
25. Kobus T, Vos PC, Hambrock T, et al. Prostate cancer aggressiveness: in vivo assessment of MR spectroscopy and diffusion-weighted imaging at 3 T. *Radiology* 2012; 265(2): 457-67.
26. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *European radiology* 2012; 22(4): 746-57.
27. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Medical decision making : an international journal of the Society for Medical Decision Making* 1991; 11(2): 88-94.


PART ONE
Chapters 2, 3, 4
Evidence acquisition

2

Accuracy of mp-MRI for prostate cancer detection: a meta-analysis

Published as

de Rooij M, Hamoen EHJ, Fütterer JJ, Barentsz JO, Rovers MM

Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis

American Journal of Roentgenology, 2014 Feb;202(2):343-51

ABSTRACT

Objective

The purpose of this diagnostic meta-analysis was to determine the diagnostic accuracy of multiparametric magnetic resonance imaging (mp-MRI) for prostate cancer (PCa) detection using anatomical T2-weighted imaging (T2WI) combined with two functional techniques: diffusion weighted imaging (DWI) and dynamic contrast-enhanced MRI (DCE-MRI).

Materials and Methods

We searched electronic databases, including MEDLINE (PubMed), Embase, and the Cochrane Central Register of Controlled Trials (CENTRAL) up to February 3, 2012. We included diagnostic accuracy studies using a combination of T2WI, DWI and DCE-MRI to detect PCa with histopathologic data from prostatectomy or biopsy as the reference standard. The methodologic quality was assessed with version 2 of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool by two independent reviewers. Sensitivity and specificity of all studies were calculated from 2x2 tables, and the results were plotted in a hierarchical summary receiver operating characteristic plot.

Results

Seven studies that met the inclusion criteria (526 patients) could be analyzed. The pooled data showed a specificity of 0.88 (95% confidence interval (CI) 0.82-0.92) and sensitivity of 0.74 (95% CI 0.66-0.81) for PCa detection, with negative predictive values (NPVs) ranging from 0.65 to 0.94. Subgroup analysis showed no significant difference between the subgroups.

Conclusion

The high specificity with variable but high NPVs and sensitivites imply a potential role for mp-MRI in detecting PCa.

INTRODUCTION

Prostate cancer (PCa) is the most common noncutaneous cancer in men.[1] Although most types of PCa grow slowly and may need minimal or no treatment, other types are aggressive and can spread quickly. PCa that is detected early has a better chance of successful treatment. Therefore, detection of PCa in an early stage is important but remains challenging.

The currently used diagnostic tools are digital rectal examination (DRE), serum prostate specific antigen (PSA) – a non-specific blood test – and transrectal ultrasound guided biopsy (TRUS-GB) – a standardized but untargeted method.[2] Because of the limitations of these available diagnostic tools, much effort is being put into improving the accuracy of PCa detection.

Advances in magnetic resonance imaging (MRI) techniques show potential for improving the diagnostic accuracy of MRI for PCa detection. A recently developed multi-parametric MRI (mp-MRI) approach that combines anatomic T2-weighted imaging (T2WI) with functional data appears to be one of the most promising techniques for PCa detection.[3-9] The addition of functional MRI techniques can provide metabolic information, display altered cellularity, and aid in non-invasive characterization of tissue and tumour vascularity.[10] Although these techniques have not been implemented broadly in daily clinical practice yet, they are increasingly mentioned in PCa guidelines.[11] The latest diagnostic consensus statement by the European Society of Urogenital Radiology (ESUR) recommends anatomic T2WI combined with at least two functional techniques: diffusion weighted imaging (DWI), dynamic contrast-enhanced MRI (DCE-MRI), and optionally magnetic resonance spectroscopy imaging (MRSI).[2] However, he accuracy of this method has not been studied systematically. We therefore performed a systematic review and meta-analysis to determine the diagnostic accuracy of the ESUR recommendation, that is, combining T2WI with DWI and DCE-MRI for the detection of PCa.

MATERIALS AND METHODS

Data sources and searches

We systematically searched the electronic databases MEDLINE (U.S. National Library of Medicine) (PubMed), Embase (Elsevier) and the Cochrane Central Register of Controlled

Trials (CENTRAL) (The Cochrane Collaboration) to identify all relevant studies. The search strategy involved a filter combining imaging modality keywords with PCa keywords in the titles and abstracts as follows: (prostate OR PCa OR PSA OR prostatic) AND (MR OR NMR OR NMRI OR MRI OR magnetic resonance OR ADC OR DWI OR DCE OR diffusion weighted OR dynamic contrast OR multiparametric). To retrieve additional publications, we manually searched reference lists from the included articles and relevant systematic and narrative reviews on the topic. No restrictions on language or date were used in this comprehensive search. The last search was performed on February 3, 2012. We imported all citations identified by the MEDLINE and Embase search strategies into a bibliographic database of EndNote, version X5 (Thomas Reuters, New York City, NY).

Study selection

We screened all retrieved articles and included studies when they compared T2WI and the functional MR techniques DWI and DCE-MRI with histologic results from prostate biopsies or prostatectomy specimens as the reference standard in patients with suspected or previously diagnosed PCa. One reviewer (MdR) performed the first screening of titles and abstracts to select eligible studies. Subsequently, two reviewers (MdR, EH) independently assessed the eligibility by reading the articles.

Data extraction and quality assessment

To obtain 2 x 2 contingency tables from the included studies, we extracted or calculated true-negative (TN), false-negative (FN), true-positive (TP), and false-positive (FP) results of mp-MRI for the detection of PCa. A standardized form was used to extract additional data on patient characteristics, imaging protocols, and methodological characteristics. The authors of the studies that did not report all sufficient data were asked to provide additional information.

We extracted the following data: patient age, PSA level, Gleason score, previous prostate biopsies, cancer status (suspected or detected), MR imager model and manufacturer, magnetic field strength (in Teslas), use of an endorectal coil, use of other coils, T2WI sequences, DWI acquisition parameters, DCE-MRI acquisition parameters, use of additional techniques, year of publication, study population, reference standard (prostate biopsy or prostatectomy specimens), patient enrolment, study design, blinding, number of readers, region of interest, and scoring system of each modality and of the combination of modalities.

Quality assessment of the included studies was performed by two independent reviewers using the recently developed version 2 of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for diagnostic accuracy studies.[12] Any disagreements were discussed and resolved by consensus.

Data synthesis and analysis

We constructed 2×2 contingency tables and calculated sensitivity and specificity with 95% confidence intervals (CI) for each study individually. We drew forest plots to show variation and to explore heterogeneity for sensitivity and specificity and plotted their results on a receiver operating characteristic (ROC) plot. We used the Metadas tool within the statistical software package SAS to carry out the meta-analyses.[13] The analyses were imported to RevMan5 (The Nordic Cochrane Center, Copenhagen, Denmark) and used to fit the hierarchical summary receiver operating characteristic (HSROC) plot. Because of the substantial heterogeneity of the included studies, we analyzed subgroups of three clinically relevant covariates; reference test (prostatectomy and biopsy), method of analysis (region-based and patient-based), and localization of the tumour (peripheral zone and whole gland).

RESULTS

Literature search

The systematic literature search identified 10,166 records. Screening the titles and abstracts and removing duplicates yielded 367 potentially eligible studies using mp-MRI techniques. Another 241 studies were excluded because they did not determine diagnostic accuracy, leaving 126 studies for formal evaluation (see also Figure 1). Seven studies used T2WI, DWI, and DCE-MRI for the detection of PCa and were included in the meta-analysis.[14-20] Other studies were not included because they did not use T2WI, DWI, and DCE-MRI as a combination of mp-MRI techniques (*n*=103) or did not report a combined analysis of these techniques or available data were insufficient to construct a 2 x 2 contingency table (*n*=16). Manual searching reference lists of narrative and systematic reviews, position papers, and editorials did not yield any additional results.

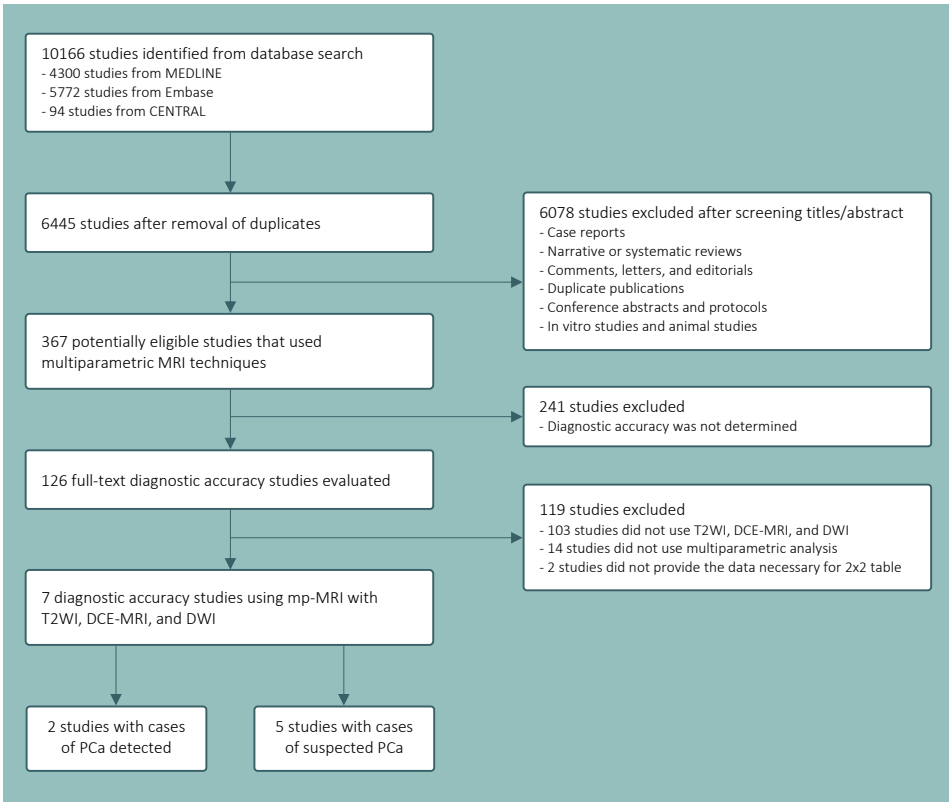


Figure 1 Flow of studies through the selection process.

Study description

Patient characteristics, technical parameters, and study design of the seven included studies are presented in Table 1. In total, 526 patients were evaluated, with a median age ranging from 63 to 69 years, a median serum PSA level varying from 7 to 11.1 ng/mL, and a Gleason score ranging from 4 to 10.

Assessment of study quality

All studies were evaluated on their methodologic quality by two independent reviewers using the QUADAS-2 tool (see also Figure 2, 3 and Supplemental Figure 1).[12] The quality of the seven studies varied. Risk of bias regarding patient selection was low in three of the seven studies,[16,17,20] whereas four studies had a high risk of bias for patient selection.[14,15,18,19] The high risk was caused by the unavailability of data on patient enrolment and inappropriate exclusion. The risk of bias regarding the index test was low in five studies[14-18] and high in one study.[19] For the remaining study, the risk of bias was unclear because information

about blinding was not sufficient.[20] The risk of bias regarding the reference standard was low in two studies[14,20] and high in five studies[15-19] because these latter studies used TRUS-GB or transperineal biopsy as the reference standard instead of a prostatectomy specimen. Furthermore, two of the five studies using TRUS-GB did not report sufficient information about the protocols and number of cores taken.[15,20] The five studies using biopsy as the reference standard comprise studies with targeted biopsies, which have a lower risk of bias than random biopsies. This factor is taken into account in assessing concerns regarding applicability, where these studies are assigned as having a low risk of bias (see also Figure 3). Risk of bias regarding flow and timing was low in four studies[16,17,19,20] and high in three studies[14,15,18] because these studies did not include all patients in the final analysis.

Diagnostic accuracy of a combined analysis of T2WI/DWI/DCE-MRI

In total, seven studies including 526 patients were considered in the final analysis. For each study, the number of TPs, FPs, FNs, and TNs are shown in Figure 5. Pooled sensitivity and specificity values for all studies were 0.74 (95% CI 0.66-0.81) and 0.88 (95% CI 0.82-0.92), respectively. Negative predictive values (NPVs) were high, varying from 0.65 to 0.94, and positive predictive values (PPVs) – with larger variability – ranged from 0.31 to 0.95 (see also Table 2). Figure 4 shows the HSROC plot with 95% CI area and summary point.

Clinically relevant subgroups

Regarding the reference standard, there was no significant difference in pooled sensitivity and specificity between studies using prostatectomy and studies using biopsy. The two prostatectomy studies showed a pooled sensitivity and specificity of 0.69 (95% CI 0.52-0.82) and 0.93 (95% CI 0.81-0.97), respectively. The five studies using biopsy as the reference test (TRUS-GB or transperineal biopsy) showed a pooled sensitivity and specificity of 0.76 (95% CI 0.66-0.84) and 0.86 (95% CI 0.79-0.91), respectively. In addition, Figure 6 shows forest plots of the pooled estimates overall and for different subgroups.

Table 1a Patient characteristics.

Author	Year	No. of patients	No. of patients with PCa	Age, median (year)	Age, range (year)	PSA, median (ng/mL)	PSA, range (ng/mL)	Gleason score, median	Gleason score, range	Previous negative biopsies	Cancer status
Delongchamps et al.	2011	57	57	63	54-76	7	2.8-28.0	NR	NR	NR	Detected
Iwazawa et al.	2011	178	72	68.8 (mean)	41-86	20.5 (mean)	4.0-568.5	7.04 (mean)	6-9	NR	Suspected
Tamada et al.	2011	50	35	70 (mean)	40-84	6.7 (mean)	4.1-9.9	7	6-10	NR	Suspected
Vilanova et al.	2011	70	38	63.5 (mean)	43-87	7.4	4.0-17.2	7	5-8	No	Suspected
Kitajima et al.	2010	53	30	69	56-84	11.1	4.2-112.1	NR	NR	NR	Suspected
Yoshizako et al.	2008	35	23	65	52-76	NR	NR	NR	6-9	NR	Detected
Tanimoto et al.	2007	83	44	67.4 (mean)	53-87	19.4 (mean)	4.3-332.1	6.9 (mean)	4-9	NR	Suspected

Abbreviations: NR = not reported; PCa = prostate cancer; PSA = prostate specific antigen.

Table 1b Technical characteristics.

Study	Equipment				T2WI			DWI					DCE-MRI				
	Model	Field Strength (T)	Coil	ERC	Planes	Slice mm	FOV (cm)/ matrix	Plane	Slice mm	FOV (cm)/ matrix	b-values	ADC-map	Plane	Slice mm	FOV (cm)/ matrix	Temporal resolution (s)	T1WI
Delongchamps et al.	Avanto, Siemens Medical Systems	1.5	Pelvic phased-array coil	yes	Axial, sagittal	1	18/186 x 256	Axial, sagittal	3.5	18/186 x 256	0, 800	yes	Axial	3.5	NR	8.5	yes
Iwazawa et al.	Magnetom Symphony, Siemens Medical	1.5	Eight-channel phased-array coil	no	Axial	5	23/512 x 352	Axial	4	30/128 x 114	0, 1000	yes	Axial	2.5	23/NR	NR	NR
Tamada et al.	Signa Excite High speed, GE Healthcare	1.5	Multichannel phased-array torso coil	no	Axial, coronal	4	24/320 x 320 (FSE), 24/256 x 256 (EPI)	Axial	5	36/256 x 256	0, 800	yes	Axial	5 (FSE) and 3 (LAVA)	27/256 x 128 (FSE), 35/288 x 192 (LAVA)	NR	yes
Vilanova et al.	Signa Horizon HDx, GE Healthcare	1.5	Pelvic four-channel phased-array coil	yes	Axial, coronal	3	14/256 x 192	Axial	3	26/128 x 128	0, 1000	yes	Axial	4	26/160 x 256	6	yes
Kitajima et al.	Magnetom Trio Tim, Siemens Medical Solution,	3	Pelvic multichannel phased-array body coil	no	Axial, coronal	3	20/320 x 320	Axial	3	35 x 25/128 x 92	0, 1000	yes	Axial	3	30/256 x 236	NR	yes
Yoshizako et al.	Signa CV/i version 9.1, GE Healthcare	1.5	Pelvic phased-array coil	no	Axial, coronal	5	22/256 x 224	Axial	5	42 x 21/256 x 256	0, 1000	yes	Axial	5	26 x 21/256 x 160	NR	yes
Tanimoto et al.	Signa Excite XI, GE Healthcare	1.5	Eight-channel torso-array coils	no	Axial	5	18/288 x 192	Axial	5	36/160 x 128	0, 1000	yes	Axial	5	18/256 x 160	NR	yes

Abbreviations: ADC = apparent diffusion coefficient; DCE-MRI = dynamic contrast-enhanced MRI; DWI = diffusion-weighted imaging; EPI = echo planar imaging; ERC = endorectal coil; FOV = field of view; FSE = fast spin echo; LAVA = liver imaging with volume acceleration; NR = not reported; T = Tesla; T1WI = T1-weighted imaging; T2WI = T2-weighted imaging.

Table 1c Study characteristics.

Study	Study population	Reference test	Patient enrolment	Study design	Blinding	Region of interest	Single pathologic correlative	T2WI scoring	DWI scoring	DCE-MRI scoring	Combined analysis	Readers
Delongchamps et al.	Candidates for prostatectomy	Histopathology	Consecutive	NR	yes	PZ and TZ	Region based (8 regions)	dichotomy	3-point scale with ADC thresholds	3-point scale	>5/9 is positive for PCa (TZ), >6/9 is positive for PCa (PZ)	2 (consensus)
Iwazawa et al.	Elevated PSA (>4 ng/mL)	TRUS-GB (10-core + targeted)	Consecutive	Retrospective	yes	Whole gland, PZ and TZ	Region based (4 regions)	4-point scale	4-point scale, no ADC thresholds	4-point scale	Positive for PCa when at least one of the three methods is abnormal (3 or 4)	2 (independent)
Tamada et al.	Elevated PSA (4-10 ng/mL)	TRUS-GB (12-core)	NR	Retrospective	yes	Whole gland	Patient and region based (8 regions)	dichotomy	dichotomy	dichotomy	Positive for PCa when at least one of the three methods is abnormal	2 (consensus)
Vilanova et al.	Elevated PSA (>4 ng/mL) and free to total PSA ratio <20%. No prior biopsy	TRUS-GB (systematic + targeted (n=57) or histopathology (n=13)	NR	Retrospective	yes	Whole gland	Region based (2 regions)	5-point scale	5-point scale, with ADC thresholds	5-point scale with curve type analysis	Positive for PCa when at least one of the three methods is abnormal (3, 4, or 5)	3 (consensus)
Kitajima et al.	Elevated PSA	Transperineal biopsy (20-core)	Consecutive	Retrospective	yes	Whole gland	Region based (8 regions)	5-point scale	5-point scale	5-point scale	Table with combined scoring system	2 (consensus)
Yoshizako et al.	Candidates for prostatectomy	Histopathology	Not consecutive	Retrospective	no	TZ	Tumour based	5-point scale	5-point scale	5-point scale	>8/15 is positive for PCa	2 (consensus)
Tanimoto et al.	Elevated PSA (>4.0 ng/mL)	TRUS-GB	Consecutive	Prospective	NR	Whole gland	Patient based	5-point scale	5-point scale	5-point scale	Positive for PCa when at least one of the three methods is abnormal (3, 4, or 5)	2 (consensus)

Abbreviations: ADC = apparent diffusion coefficient; DCE-MRI = dynamic contrast-enhanced MRI; DWI = diffusion weighted imaging; NR = not reported; PCa = prostate cancer; PSA = prostate specific antigen; PZ = peripheral zone; T2WI = T2-weighted imaging; TRUS-GB = transrectal ultrasound guided biopsy; TZ = transition zone.

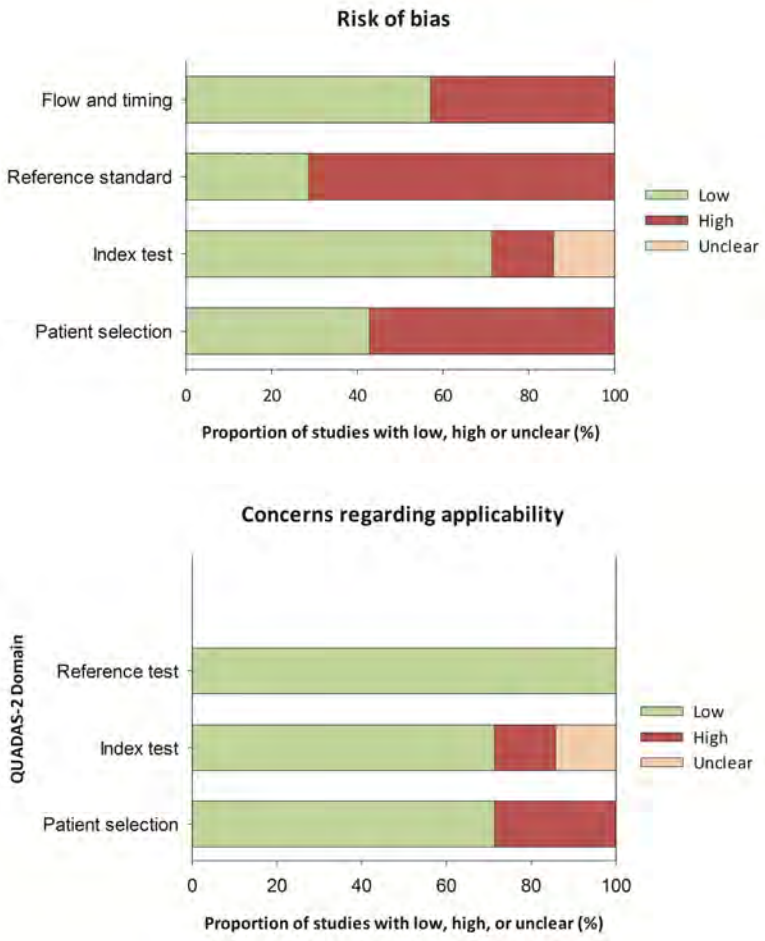


Figure 2 and 3 Methodological quality summary of risk of bias (upper) and applicability (lower) on each of the four QUADAS-2 domains.

The four studies that assessed accuracy based on classification of region level of the prostate showed a pooled sensitivity and specificity of 0.71 (95% CI 0.63-0.78) and 0.89 (95% CI 0.83-0.94), respectively. The studies reporting on a patient or tumour level could not be pooled because available data were not sufficient.

When subgroup comparisons based on the localization of the analysed tumours were made, the pooled sensitivity values were comparable in the studies analysing peripheral zone tumours and studies analysing the whole prostate gland. The pooled sensitivity and specificity of the studies analysing peripheral zone tumours were 0.81 (95% CI 0.75-0.85) and 0.91 (95% CI 0.67 0.98), respectively. The studies analysing the whole prostate showed

a pooled sensitivity of 0.78 (95% CI 0.65-0.87) and a pooled specificity of 0.88 (95% CI 0.80-0.94). We were not able to pool calculated estimates of the transition zone.

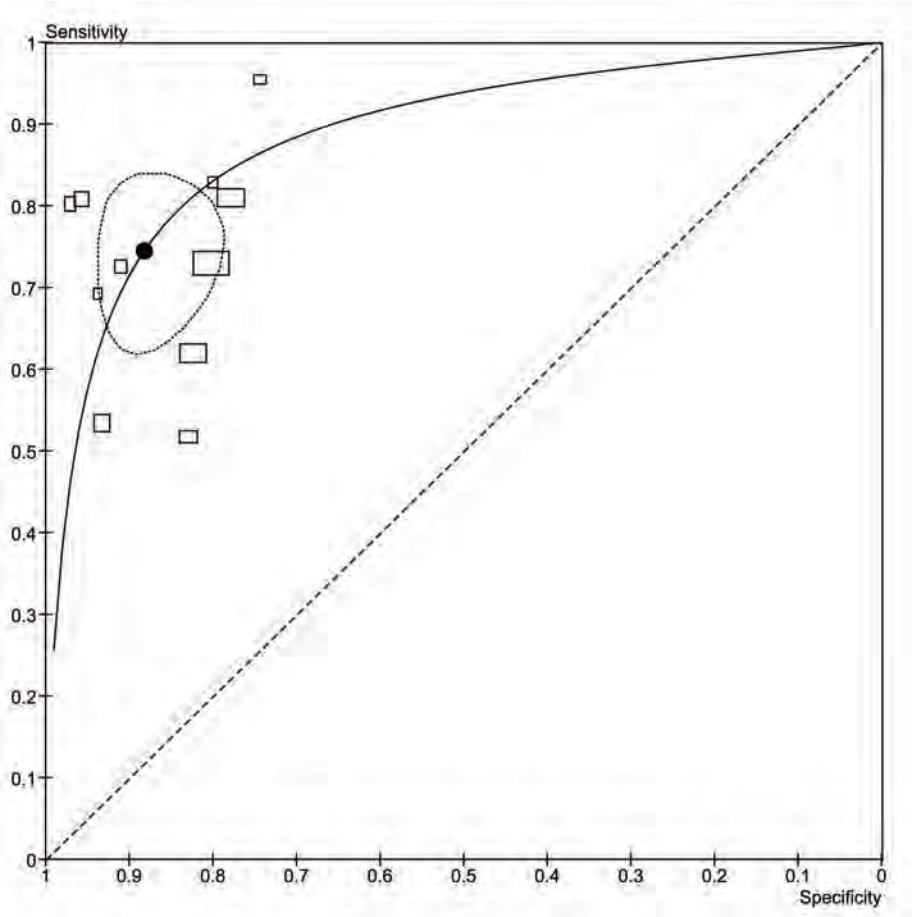


Figure 4 Hierarchical summary receiver operating characteristic (HSROC) plot with summary point and 95% confidence region.

Study	TP	FP	FN	TN	Reference	Localization	Analysis	Sensitivity	Specificity
Delongchamps 2010 PZ	73	4	18	133	Prostatectomy	Peripheral zone	Region	0.80 [0.71, 0.88]	0.97 [0.93, 0.99]
Yoshizako 2008	18	1	8	15	Prostatectomy	Transition zone	Tumor	0.69 [0.48, 0.86]	0.94 [0.70, 1.00]
Delongchamps 2010 TZ	15	34	14	165	Prostatectomy	Transition zone	Region	0.53 [0.34, 0.69]	0.83 [0.77, 0.88]
Iwazawa 2010 PZ	149	117	35	411	Biopsy	Peripheral zone	Region	0.81 [0.75, 0.86]	0.78 [0.74, 0.81]
Iwazawa 2010 TZ	83	102	51	476	Biopsy	Transition zone	Region	0.62 [0.53, 0.70]	0.82 [0.79, 0.85]
Tanimoto 2007	42	10	2	29	Biopsy	Whole gland	Patient	0.95 [0.85, 0.99]	0.74 [0.58, 0.87]
Tamada 2011 Patient	29	3	6	12	Biopsy	Whole gland	Patient	0.83 [0.66, 0.93]	0.80 [0.52, 0.96]
Vilanova 2011	37	8	14	81	Biopsy	Whole gland	Region	0.73 [0.58, 0.84]	0.91 [0.83, 0.96]
Kitajima 2010	80	14	19	311	Biopsy	Whole gland	Region	0.81 [0.72, 0.88]	0.96 [0.93, 0.98]
Iwazawa 2010 Whole gland	232	219	86	887	Biopsy	Whole gland	Region	0.73 [0.68, 0.78]	0.80 [0.78, 0.83]
Tamada 2011 Region	55	20	48	277	Biopsy	Whole gland	Region	0.53 [0.43, 0.63]	0.93 [0.90, 0.96]

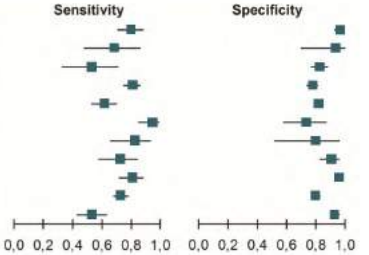


Figure 5 Forest plots of the included studies.

Study	TP	FP	FN	TN	Reference	Localization	Analysis	Sensitivity	Specificity
Delongchamps 2010 PZ	73	4	18	133	Prostatectomy	Peripheral zone	Region	0.80 [0.71, 0.88]	0.97 [0.93, 0.99]
Yoshizako 2008	18	1	8	15	Prostatectomy	Transition zone	Tumor	0.69 [0.48, 0.86]	0.94 [0.70, 1.00]
Delongchamps 2010 TZ	15	34	14	165	Prostatectomy	Transition zone	Region	0.53 [0.34, 0.69]	0.83 [0.77, 0.88]
Iwazawa 2010 PZ	149	117	35	411	Biopsy	Peripheral zone	Region	0.81 [0.75, 0.86]	0.78 [0.74, 0.81]
Iwazawa 2010 TZ	83	102	51	476	Biopsy	Transition zone	Region	0.62 [0.53, 0.70]	0.82 [0.79, 0.85]
Tanimoto 2007	42	10	2	29	Biopsy	Whole gland	Patient	0.95 [0.85, 0.99]	0.74 [0.58, 0.87]
Tamada 2011 Patient	29	3	6	12	Biopsy	Whole gland	Patient	0.83 [0.66, 0.93]	0.80 [0.52, 0.96]
Vilanova 2011	37	8	14	81	Biopsy	Whole gland	Region	0.73 [0.58, 0.84]	0.91 [0.83, 0.96]
Kitajima 2010	80	14	19	311	Biopsy	Whole gland	Region	0.81 [0.72, 0.88]	0.96 [0.93, 0.98]
Iwazawa 2010 Whole gland	232	219	86	887	Biopsy	Whole gland	Region	0.73 [0.68, 0.78]	0.80 [0.78, 0.83]
Tamada 2011 Region	55	20	48	277	Biopsy	Whole gland	Region	0.53 [0.43, 0.63]	0.93 [0.90, 0.96]

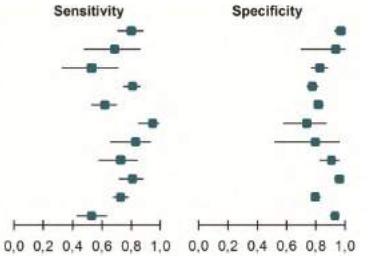


Figure 6 Forest plots of the pooled estimates; overall and different subgroups.

Table 2 Prevalence of prostate cancer and predictive values.

Study	Year	Prevalence (absolute)	Analysis	PPV	NPV
Delongchamps PZ	2010	91	Region	0.95	0.88
Yoshizako	2008	26	Tumour	0.95	0.65
Delongchamps TZ	2010	29	Region	0.31	0.92
Iwazawa PZ	2010	184	Region	0.56	0.92
Iwazawa TZ	2010	134	Region	0.45	0.90
Tanimoto	2007	44	Patient	0.81	0.94
Tamada patient based	2011	35	Patient	0.91	0.67
Vilanova	2011	51	Region	0.82	0.85
Kitajima	2010	99	Region	0.85	0.94
Iwazawa whole gland	2010	318	Region	0.51	0.91
Tamada region based	2011	103	Region	0.73	0.85

Abbreviations: NPV = negative predictive value; PPV = positive predictive value; PZ = peripheral zone; TZ = transition zone.

DISCUSSION

The results of this diagnostic meta-analysis on the accuracy of mp-MRI for PCa detection using the combination of T2WI, DWI, and DCE-MRI revealed a high overall sensitivity and specificity. The overall methodologic quality of the included studies was fair, but large heterogeneity was reported. Nevertheless, subgroup analyses did not show considerable differences among various subgroups.

To date, most studies have reported various sensitivity and specificity values for the accuracy of anatomic T2WI with or without one or more additional functional techniques for the detection of PCa. Recently, a systematic review and meta-analysis was published on the diagnostic accuracy of T2WI combined with DWI compared with T2WI alone.[21] The meta-analysis of the 10 included studies showed a higher diagnostic accuracy for T2WI combined with DWI (sensitivity and specificity of 0.72 and 0.81, respectively) than for T2WI alone (0.62 and 0.77).

The major strength of this diagnostic meta-analysis is that this is the first meta-analysis to investigate the accuracy of the combination of anatomic T2WI and two functional techniques; DWI and DCE-MRI, as recommended by the ESUR guidelines.[2] Furthermore, we

are one of the first groups to undertake a meta-analysis of MRI using HSROC methods that are available using the Metadas macro for SAS.[13]

This diagnostic meta-analysis showed that the accuracy of mp-MRI shows potential for the detection of PCa. Although the FN rate of 26% still might be too high, TRUS-GB tends to miss tumours as well, with detection rates of 10-19% on repeat TRUS-GB[22,23] and up to 59% on MRGB after two negative TRUS-GB sessions.[24] However, whether those FNs are clinically significant or insignificant tumours is still open to debate. A future randomized multi-centre diagnostic trial comparing TRUS-GB with mp-MRI is needed to study potential benefits, harms, and cost-effectiveness in more detail.

The recommendation of the ESUR of using T2WI, DWI, and DCE-MRI for PCa detection is based on expert opinion, and the question remains about whether this strategy is the best multiparametric combination because prospective validation studies have not yet been performed. Of the seven included studies, four studies[16-19] recommend using both DWI and DCE-MRI as additional techniques and show significant differences in performance compared with the use of DWI or DCE-MRI alone. The other three studies[14,15,20] show no change in performance or worse results when comparing the combination of T2WI, DWI, and DCE-MRI with T2WI and DWI.

Some potential limitations should be mentioned. First, considerable heterogeneity was identified among the included studies, with differences in reference tests, prostate regions, patient characteristics, and method of analysis. Four of seven studies used TRUS-GB, as the reference standard for the primary diagnosis of PCa. However, this technique mainly randomly samples the posterior part of the gland, tends to miss tumours on first systematic biopsy[22], and has been reported to underestimate Gleason grade in 43% of the cases.[25] The studies also analysed different regions of the prostate and patients with other baseline characteristics, such as PSA level. Two studies included patients with mean PSA values of 20.51 and 19.4 ng/mL; these mean PSA values are considerably higher than levels observed in cases of early phase PCa.[17,20] In these studies a large tumour burden can be expected, with subsequently a possible higher accuracy of mp-MRI. The method of analysis varied among the different studies, with studies using a per-patient or region-based approach. For the studies that used a region-based approach, the prostate was subdivided in a varying number of regions, ranging from two to eight, which can artificially increase the specificity by generating more TN regions. Despite this heterogeneity, the pooled estimates of the different subgroups were comparable with the overall summary values.

Second, despite the small number and heterogeneity of the studies, we decided to pool the results because sensitivity and subgroup analyses showed no significant differences. We believe that there are many good examples of reviews, such as Cochrane reviews, performing a meta-analysis of only a few studies that have been proven useful. Furthermore, this review can be considered as a starting point and can be updated as soon as new evidence becomes available.

Third, the overall estimated sensitivity and specificity values were based on all included studies. Although some studies reported data on peripheral zone, transition zone, and whole gland separately and some reported the results of patient- and region-based analysis, these studies were not equally represented in this analysis.

Fourth, the technical parameters, such as the use of an endorectal coil, field strength, and b-values, were not similar among the included studies. Furthermore, the analysis of DCE-MRI could be profoundly influenced by the difference in spatial resolution and temporal resolution and the pharmacokinetic model that is used. The recently developed ESUR guidelines can be used to fulfil the minimal technical and image acquisition parameter requirements for an MRI protocol to detect PCa.[2]

Finally, scoring systems for reporting the prostate images were not similar in all studies. Several studies used dichotomized scoring to distinguish between normal and abnormal appearance, whereas other studies used 3- or 5-point Likert scales for the overall or separate imaging techniques. ESUR experts developed a structured reporting system (the Prostate Imaging Reporting and Data System, or PI-RADS) with a standardized subscore for each sequence (T2WI, DWI, and DCE-MRI), the subscores were summarized in a final score that could range between 1 and 5,[2] similar to the standardized score used for breast MRI (BI-RADS). Although this scoring system has not yet been validated, the ESUR classification, which was recently adopted by the American College of Radiology, is the best available guideline for using mp-MRI in the diagnosis of PCa.

CONCLUSIONS

The high specificity with variable but high NPVs and sensitivities imply a possible role for mp-MRI before biopsy in detecting PCa.

REFERENCES

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60: 277-300.
2. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol* 2012; 22: 746-57.
3. Sciarra A, Panebianco V, Salciccia S, et al. Modern role of magnetic resonance and spectroscopy in the imaging of prostate cancer. *Urol Oncol* 2011; 29: 12-20.
4. Puech P, Sufana Iancu A, Renard B, Villers A, Lemaitre L. Detecting prostate cancer with MRI - why and how. *Diagn Interv Imaging* 2012; 93: 268-78.
5. Villers A, Marliere F, Ouzzane A, Puech P, Lemaitre L. MRI in addition to or as a substitute for prostate biopsy: The clinician's point of view. *Diagn Interv Imaging* 2012; 93: 262-7.
6. Engelbrecht MR, Puech P, Colin P, Akin O, Lemaitre L, Villers A. Multimodality magnetic resonance imaging of prostate cancer. *J Endourol* 2010; 24: 677-84.
7. Mazaheri Y, Shukla-Dave A, Muellner A, Hricak H. MRI of the prostate: clinical relevance and emerging applications. *J Magn Reson Imaging* 2011; 33: 258-74.
8. Hoeks CM, Barentsz JO, Hambrock T, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology* 2011; 261: 46-66.
9. Sciarra A, Barentsz JO, Bjartell A, et al. Advances in magnetic resonance imaging: how they are changing the management of prostate cancer. *Eur Urol* 2011; 59: 962-77.
10. Pinto F, Totaro A, Calarco A, et al. Imaging in prostate cancer diagnosis: present role and future perspectives. *Urol Int* 2011; 86: 373-82.
11. Heidenreich A, Bastian PJ, Bellmunt J. Guidelines on prostate cancer. European Association of Urology; 2012.
12. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529-36.
13. Takwoingi Y, Deeks, J. METADAS: A SAS macro for meta-analysis of diagnostic accuracy studies. v1.3 ed; 2008.
14. Delongchamps NB, Rouanne M, Flam T, et al. Multiparametric magnetic resonance imaging for the detection and localization of prostate cancer: combination of T2-weighted, dynamic contrast-enhanced and diffusion-weighted imaging. *BJU Int* 2011; 107: 1411-8.
15. Vilanova JC, Barcelo-Vidal C, Comet J, et al. Usefulness of prebiopsy multifunctional and morphologic MRI combined with free-to-total prostate-specific antigen ratio in the detection of prostate cancer. *AJR American journal of roentgenology* 2011; 196: W715-22.
16. Kitajima K, Kaji Y, Fukabori Y, Yoshida K, Suganuma N, Sugimura K. Prostate cancer detection with 3 T MRI: comparison of diffusion-weighted imaging and dynamic contrast-enhanced MRI in combination with T2-weighted imaging. *J Magn Reson Imaging* 2010; 31: 625-31.
17. Iwazawa J, Mitani T, Sassa S, Ohue S. Prostate cancer detection with MRI: is dynamic contrast-enhanced imaging necessary in addition to diffusion-weighted imaging? *Diagnostic and interventional radiology (Ankara, Turkey)* 2011; 17: 243-8.
18. Tamada T, Sone T, Higashi H, et al. Prostate cancer detection in patients with total serum prostate-specific antigen levels of 4-10 ng/mL: diagnostic efficacy of diffusion-weighted imaging, dynamic contrast-enhanced MRI, and T2-weighted imaging. *AJR American journal of roentgenology* 2011; 197: 664-70.
19. Yoshizako T, Wada A, Hayashi T, et al. Usefulness of diffusion-weighted imaging and dynamic contrast-enhanced magnetic resonance imaging in the diagnosis of prostate transition-zone cancer. *Acta radiologica (Stockholm, Sweden : 1987)* 2008; 49: 1207-13.

20. Tanimoto A, Nakashima J, Kohno H, Shinmoto H, Kuribayashi S. Prostate cancer screening: the clinical value of diffusion-weighted imaging and dynamic MR imaging in combination with T2-weighted imaging. *J Magn Reson Imaging* 2007; 25: 146-52.

21. Wu LM, Xu JR, Ye YQ, Lu Q, Hu JN. The clinical value of diffusion-weighted imaging in combination with T2-weighted imaging in diagnosing prostate carcinoma: a systematic review and meta-analysis. *AJR American journal of roentgenology* 2012; 199(1): 103-10.

22. Djavan B, Ravery V, Zlotta A, et al. Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop? *J Urol* 2001; 166: 1679-83.

23. Gosselaar C, Roobol MJ, Roemeling S, Wolters T, van Leenders GJ, Schroder FH. The value of an additional hypoechoic lesion-directed biopsy core for detecting prostate cancer. *BJU Int* 2008; 101: 685-90.

24. Hambrock T, Hoeks C, Hulsbergen-van de Kaa C, et al. Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort. *Eur Urol* 2012; 61: 177-84.

25. Scattoni V, Zlotta A, Montironi R, Schulman C, Rigatti P, Montorsi F. Extended and saturation prostatic biopsy in the diagnosis and characterisation of prostate cancer: a critical analysis of the literature. *Eur Urol* 2007; 52: 1309-22.

3

Use of PI-RADS for prostate cancer detection with mp-MRI: a diagnostic meta-analysis

Published as

Hamoen EHJ, de Rooij M, Witjes JA, Barentsz JO, Rovers MM

Use of the Prostate Imaging Reporting and Data System (PI-RADS) for prostate cancer detection with multiparametric magnetic resonance imaging: a diagnosticmeta-analysis

European Urology, 2015 Jun;67(6):1112-21

ABSTRACT

Context

In 2012, an expert panel of the European Society of Urogenital Radiology (ESUR) published the Prostate Imaging Reporting and Data System (PI-RADS) for prostate cancer (PCa) detection with multiparametric magnetic resonance imaging (mp-MRI). Since then, many centres have reported their experiences.

Purpose

To review the diagnostic accuracy of PI-RADS for PCa detection with mp-MRI.

Evidence acquisition

We searched Medline and Embase up to March 20, 2014. We included diagnostic accuracy studies since 2012 that used PI-RADS with mp-MRI for PCa detection in men, using prostatectomy or biopsy as the reference standard. The methodologic quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool by two independent reviewers. Data necessary to complete 2x2 contingency tables were obtained from the included studies, and test characteristics including sensitivity and specificity were calculated. Results were pooled and plotted in a summary receiver operating characteristics plot.

Evidence synthesis

Fourteen studies (1785 patients) could be analysed. The pooled data showed sensitivity of 0.78 (95% confidence interval (CI) 0.70–0.84) and specificity of 0.79 (95% CI 0.68–0.86) for PCa detection, with negative predictive values ranging from 0.58 to 0.95. Sensitivity analysis revealed pooled sensitivity of 0.82 (95% CI 0.72–0.89) and specificity of 0.82 (95% CI 0.67–0.92) in studies with correct use of PI-RADS (ie, clear description in the methodology and no adjustment of criteria). Studies with a less strict or adjusted use of PI-RADS criteria, or unclear description of the methodology, had pooled sensitivity of 0.73 (95% CI 0.62–0.82) and specificity of 0.75 (95% CI 0.61–0.84).

Conclusions

In patients for whom PCa is suspected, PI-RADS appears to have good diagnostic accuracy in PCa detection, but no recommendation regarding the best threshold can be provided because of heterogeneity.

Patient summary

Pooling of results from all previous studies that used a relatively new 5-point scoring system for prostate-MRI showed that this scoring system appears to be able to detect prostate cancer accurately.

INTRODUCTION

Prostate cancer (PCa) is the most common noncutaneous malignancy in men in Western countries; in 2011, 903,500 new cases and 258,400 deaths were recorded worldwide.[1] Multiparametric magnetic resonance imaging (mp-MRI) is increasingly being used for PCa diagnosis because of its growing availability, advances that combine anatomic and functional data, and the increasing number of studies confirming its diagnostic reliability for PCa detection.[2-7] However, mp-MRI is not utilized in daily clinical practice everywhere. Widespread acceptance of mp-MRI of the prostate was hampered by a lack of standard diagnostic criteria for reporting of results. This lack of a standardized reporting method led to substantial variability in interpretation. Radiologists used different types of Likert scales to characterize their level of suspicion of the presence of PCa, which was generally based on the overall impression of the radiologist.[8] To standardize the evaluation and reporting of prostate MRI, the European Society of Urogenital Radiology (ESUR) published guidelines based on expert consensus in 2012, termed the Prostate Imaging Reporting and Data System (PI-RADS).[9] These guidelines provide explicit criteria for the single-modality scores (Supplemental Table 1) and are the first attempt to standardize prostate MRI. Since then, several studies using PI-RADS with mp-MRI have been published. However, the accuracy of this scoring system has not been studied systematically. Our aim was therefore to assess the diagnostic accuracy of mp-MRI using PI-RADS as a reporting system for PCa detection in men, using prostatectomy or biopsy as the reference standard.

EVIDENCE ACQUISITION

Literature search

We systematically searched MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials from inception up to March 20, 2014 for studies evaluating the diagnostic value of PI-RADS in diagnosing PCa. The search query combined synonyms for prostate cancer with synonyms for magnetic resonance imaging and synonyms for PI-RADS or scoring system (see Appendix I for the complete search strategy). We also performed a reference and related article search. Duplicate articles were manually filtered using the bibliographic EndNote database, version X5 (Thomas Reuters, New York City, NY). Searches were restricted to publications in English.

Study selection

We included studies since 2012 if (1) the PI-RADS classification in reporting prostate MRI was used, (2) reconstruction of 2×2 tables were possible for the PI-RADS score at specified cut-off points, and (3) mp-MRI, consisting of at least T2 weighted imaging (T2WI) and two functional MR techniques, was used. Studies that focused on staging PCa patients or that used another scoring system were excluded.

Two reviewers (EH and MdR) independently assessed the eligibility of the identified papers. Any disagreements were resolved by discussion with a third reviewer (MMR).

Data extraction and quality assessment

We reviewed the included studies in duplicate and extracted the study population, number of patients included, study design, year of publication, patient age, PSA level and Gleason score for patients included, previous prostate biopsies, cancer status, and inclusion and exclusion criteria. The number of observers evaluating the images or histopathology and their experience were documented, and whether there was a form of consensus reading. Moreover, the time between MRI and pathology, the type of reference standard, and the presence of malignancy and Gleason score at pathology were extracted. Similarly, we extracted basic technical characteristics for MRI, including MRI manufacturer and model, magnetic field strength in Tesla, characteristics of sequences used (T2-weighted sequences (T2WI), diffusion weighted imaging (DWI), dynamic contrast-enhanced MRI (DCE-MRI), or magnetic resonance spectroscopy (MRS)), and type of coil used.

We assessed the risk of bias and the applicability at study cohort level using the validated Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) scoring system.[10] This is a validated tool for assessment of methodologic quality and applicability of diagnostic accuracy studies. Four domains are scored: (1) patient selection, which describes the method for patient selection and the patients included; (2) index test, which describes the test being studied and how it was conducted and interpreted; (3) reference standard, which describes the reference standard used and how it was conducted and interpreted; and (4) flow and timing, which describes the flow of patient inclusion and exclusion and the interval between the index test and the reference standard. The quality assessment was performed by two independent reviewers (EH and MdR) using a data extraction form to collect details from selected studies. Any disagreements were resolved by discussion with a third reviewer (MMR).

Data synthesis and analysis

Data from each study was summarized in 2x2 tables of true positive, false positive, true negative, and false negative values to calculate sensitivity and specificity values. Authors of studies that did not report all sufficient data were asked to provide additional information. To graphically display the sensitivity and specificity measurements at study level, we used Review Manager 5 software from the Cochrane collaboration. We drew forest plots to show variation and to explore heterogeneity for sensitivity and specificity, and plotted their results on a receiver operating characteristic (ROC) curve. Primary outcomes are pooled estimates of sensitivity and specificity with 95% confidence intervals (CI). We used the Metadas tool within the statistical software package SAS (version 9.2, SAS Institute) to carry out the meta-analyses. The analyses were imported to RevMan5 (The Nordic Cochrane Center, Copenhagen, Denmark) and used to fit the hierarchic summary ROC plot.

Publication bias was studied using Deeks funnel plots.[11] This analysis was performed using the R statistical package system (R Foundation for Statistical Computing, Vienna, Austria) with a significance level set at $p<0.05$.

To study the heterogeneity in more detail, we performed sensitivity analyses of several clinically relevant covariates: analysis method (region-based versus patient-based), lesion localization (peripheral zone, transition zone, and whole gland), outcome measure (any cancer versus significant cancer), and previous biopsies (patients with versus patients without prior biopsies versus a mixed population). The definition used to determine clinically significant disease was the one used by individual studies included in the review. Furthermore, we performed a sensitivity analysis comparing studies with a reference standard with a high risk of bias versus a low risk of bias. We considered studies that used either a combination of systematic and targeted biopsies or radical prostatectomy as likely to identify PCa. Such studies were therefore considered as having a low risk of bias. Finally, we compared studies with high concerns regarding application of PI-RADS versus studies with low applicability concerns. High concerns were present when PI-RADS criteria were adjusted, or if the application of PI-RADS was unclear or not exactly specified.

EVIDENCE SYNTHESIS

Literature search and study selection

Our search yielded 1498 unique publications (see also Figure 1). Of these publications, 192 full-text articles were reviewed for eligibility, of which 178 were excluded. Studies were excluded because they did only report on PCa staging (n=52), were review articles (n=7), or were irrelevant to the review question (n=4). In addition, 109 articles on PCa detection did not use the PI-RADS classification, and six studies using PI-RADS did not report on diagnostic accuracy or provided insufficient data for 2x2 tables. Finally, 14 publications were included, covering a total of 1785 patients.[12-25]

Study characteristics

Table 1 summarizes the characteristics of the patients included in the different studies. The median age in the studies ranged from 62 to 65 years, median PSA from 5.3-10 ng/mL, and the Gleason score from 6 to 10. Patients were included consecutively in ten studies,[13-16,18-22,24]; in the other four studies this detail was not explicitly mentioned.[12,17,23,25] The percentage of patients included with PCa varied from 33% to 100%. For 394 of the 1785 patients, no data regarding previous biopsies were reported; 84 had previous positive biopsies, 690 had no previous biopsies, and 617 had at least one negative biopsy.

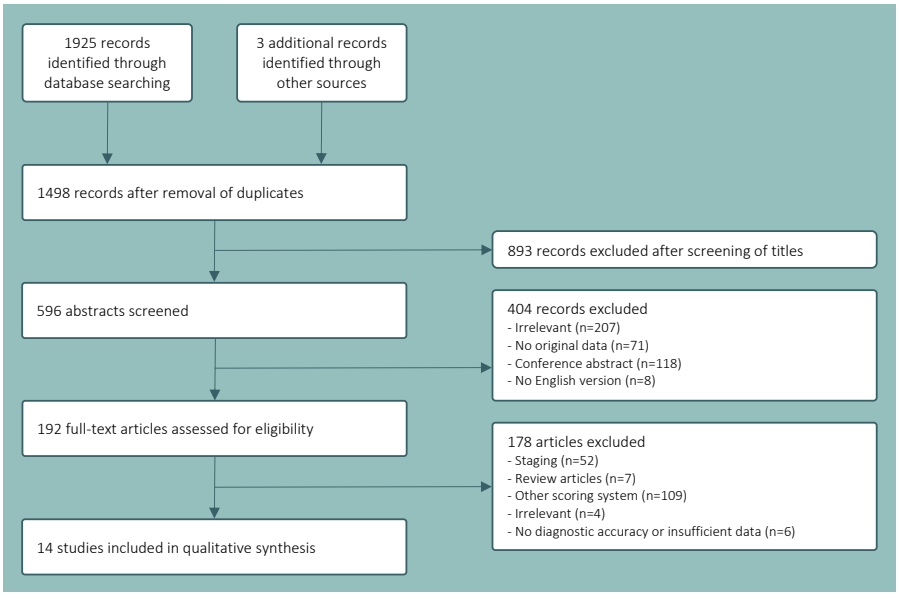


Figure 1 Flow diagram of study inclusion.

Table 1 Patient characteristics.

Study number	First author (year)	Patients (n)	Patients with PCa (n)	Age (yr)		PSA (ng/ml)		Gleason score		Previous biopsy sessions (n)	Patients diagnosed with PCa before inclusion
				Median	Range	Median	Range	Median	Range		
1	Abd-Alazeez (2014)[13]	129	92	62	41-82	5.8	1.2 – 20	6	6 – 9	0	No
2	Abd-Alazeez (2014)[12]	54	34	64	39 – 75	10	2 – 23	7	6 – 7	1 – 3	No
3	Baur (2014)[25]	55	18	66 [a]	54 – 78	10 [a]	2.9 – 65.2	7	6 – 10	0 – 6	No
4	Fiard (2013)[14]	30	14	64	61 – 67	6.3	5.2 – 8.8	7	6 – 8	0 – 2	No
5	Habchi (2013)[15]	288	152	61.8 [a]	NR	8.3 [a]	NR	NR	NR	0 – 5	Mixed
6	Junker (2013)[16]	143	39	62 [a]	NR	6.4 [a]	NR	NR	NR	≥ 1	No
7	Komai (2013)[17]	324	128	64 [a]	40 – 79	6.8 [a]	2.8 – 20	7	NR	NR	No
8	Pokorny (2014)[24]	223	142	63	NR	5.3	NR	7	6 – 10	0	No
9	Portalez (2012)[18]	129	62	64.7 [a]	47 – 79	9.6 [a]	2.7 – 40	NR	NR	1 – 4	No
10	Quentin (2013)[19]	59	28	65	52 – 83	8	4 – 49	7	6 – 9	≥ 1	No
11	Roethke (2014)[20]	64	27	64.5	49 – 77	8.3	1.3 – 22	7	NR	≥ 0	No
12	Rosenkrantz (2013)[21]	70	70	59 [a]	NR	5.9 [a]	1.9 – 18	6	6 – 9	NR	Yes
13	Schimmoeller (2013)[22]	67	28	NR	NR	NR	NR	7	6 – 9	≥ 1	No
14	Thompson (2014)[23]	150	51	62.4	NR	5.6	NR	NR	NR	≥ 0	No

[a] mean. Abbreviations: NR = not reported; PCa = prostate cancer; PSA = prostate specific antigen.

The technical characteristics of the studies included are shown in Supplemental Table 2. There were eight studies in which MRI examinations were carried out with a 3.0 Tesla scanner,[14-16,19-22,24] three studies that used a 1.5 Tesla device,[17,18,25] and three studies that alternately used 1.5 and 3.0 Tesla instruments.[12,13,23] All studies used mp-MRI consisting of T2WI, DWI, and DCE-MRI. One study reported that DCE-MRI was not used in a subset of patients.[17] MRS was used in only one study.[20] Furthermore, an endorectal coil was used in two studies.[18,25]

Table 2 describes the characteristics of the studies. All specimens were histopathologically examined; only one study specified that histopathology was interpreted without knowledge of the MRI images.[24] Most used was a combination of random biopsies and targeted cores based on MRI and ultrasound fusion (n=9).[12-18,23,24] Three studies performed only MR-guided biopsies (MRGB),[19,22,25] one study performed targeted MR-transrectal ultrasound fusion biopsies only,[20] and one used radical prostatectomy as reference standard.[21] The number of readers who evaluated prostate MRI varied from one to 13. Readers were mostly blinded, but one study did not report on blinding of the radiologists[18] and in one study readers knew that patients had undergone radical prostatectomy but were unaware of other histopathologic details.[21] Consensus reading was performed in five studies.[16,19,20,24,25] Six studies used an overall PI-RADS classification of 1-5[12,13,15,17,23,24] and the other eight studies used a sum score.[14,16,18-22,25]

The interval between MRI and histopathology was not specified in eight of the studies.[12,18-20,22-24] In most of the other studies, histopathology was obtained within 3 months after MRI.[14-17] However, one paper reported a range up to 155 days without further explanation.[21]

Only three studies reported the location of the tumours found. Rosenkrantz et al. reported 223/279 (80%) positive regions in the peripheral zone, 56/279 (20%) in the transition zone.[21] Baur et al. found PCa in 30 lesions, of which 14 (47%) were in the peripheral zone and 16 (53%) in the central gland.[25] In the study of Junker et al., PCa was present in 39 cases. Of these 39 cases, of which 22 tumours (56%) were anterior and 17 were posterior (44%). Otherwise, 17 were present in the transition zone (44%), and 22 in the peripheral zone (56%).[16]

Table 2 Study characteristics. [a]Study number as listed in Table 1; [b]Scored retrospectively.

Study number [a]	Study design	Patient enrolment	Reference test	Blinding	Type of analysis	Application of PI-RADS	Cut-off value	Readers per study cohort (n)	Outcome (any or significant cancer)
1	Prospective	Consecutive	TTB + targeted TRUS-GB	NR	Lobe	Overall 5-point scale	4	8	Any + significant
2	Prospective	NR	TTB + targeted TRUS-GB	NR	Lobe	Overall 5-point scale	4	5	Any + significant
3	Retrospective	NR	MRGB	NR	Lesion, region	Sum score	10	2	Any
4	Prospective	Consecutive	systematic TRUS-GB + MR-TRUS biopsy	NR	Patient	Sum score	NR	1	Any
5	Prospective	Consecutive	Systematic TRUS-GB + MR-TRUS (cognitive) fusion biopsy	No	Lobe, patient	Overall 5-point scale	4	13	Any
6	Retrospective	Consecutive	systematic TRUS-GB + MR-TRUS fusion biopsy	NR	Patient	Sum score	10	2	Any
7	Prospective	NR	Transperineal biopsies + systematic TRUS-GB	Yes	Patient	Overall 5-point scale. Adjusted PI-RADS (overall score deriving unclear)	4	1	Any + significant
8	Prospective	Consecutive	Systematic TRUS-GB + MRGB	NR	Patient	Overall 5-point score (derivation unclear)	3	3	Significant
9	Prospective [b]	Consecutive	Systematic TRUS-GB + MR-TRUS fusion biopsy	NR	Core	Sum score	9	NR	Any
10	Prospective	Consecutive	MRGB	NR	Lesion	Sum score, adjusted PI-RADS	10	2	Any
11	Prospective	Consecutive	MR-TRUS fusion biopsy	NR	Patient	Sum score	10	2	Any + significant
12	Retrospective	Consecutive	Radical prostatectomy	Yes (but knew of RP)	Patient, region	Sum score	8	2	Any
13	Prospective	Consecutive	MRGB	NR	Lesion	Sum score, adjusted PI-RADS	10	3	Any
14	Prospective	NR	TTB + MR-TRUS fusion biopsy	No	Patient	Overall 5-point scale (mean of 3 single-modality scores)	3	2	Significant

Abbreviations: MRGB = Magnetic Resonance Imaging guided biopsies; MR-TRUS = fusion of MRI and TRUS; NR = not reported; PI-RADS = Prostate Imaging Reporting and Data System; RP = radical prostatectomy; TTB= transperineal template biopsy; TRUS-GB = Transrectal ultrasound guided biopsy

Quality assessment

Overall, the quality of the included studies was moderate (see also Figure 2 and 3). For the patient selection domain, five studies[12,13,15,17,19] had a high risk of bias because of inappropriate inclusion or exclusion criteria. Four studies[15,17,21,25] were assigned high concerns regarding applicability. One study included a mix of patients with previous positive and negative biopsies, but the results for the two categories were not separated.[15] Two studies had a retrospective study design[21,25] and one study had unclear inclusion criteria that changed during the study, excluding patients relevant to our research question.[17] For the index test' domain, five studies[14,18,20,21,25] had a high risk of bias. Four of the five studies[14,18,20,25] did not specify the threshold used, and in one study radiologists were aware that patients had undergone radical prostatectomy.[21] For seven studies there were high concerns regarding the applicability of PI-RADS[12-15,23-25] because the overall or single-modality PI-RADS score synthesis was not specified or adjusted PI-RADS criteria were used. For the reference standard domain four studies[19,20,22,25] had a high risk of bias. We considered a combination of systematic and targeted biopsies, or radical prostatectomy, as likely to identify PCa, so studies using such a combination were considered as having a low risk of bias. Four studies[12,13,23,24] had high concerns regarding applicability because detection of significant PCa alone was the outcome measure. For the flow and timing domain, three studies had a high risk of bias[12,13,23] because the same reference standard was not applied for all the patients,[12] there was a long interval between the index test and the reference standard,[13] or not all patients were included in the analysis.[23]

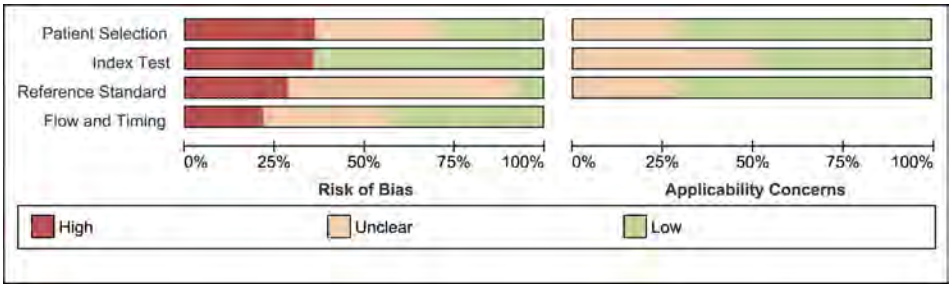


Figure 2 Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

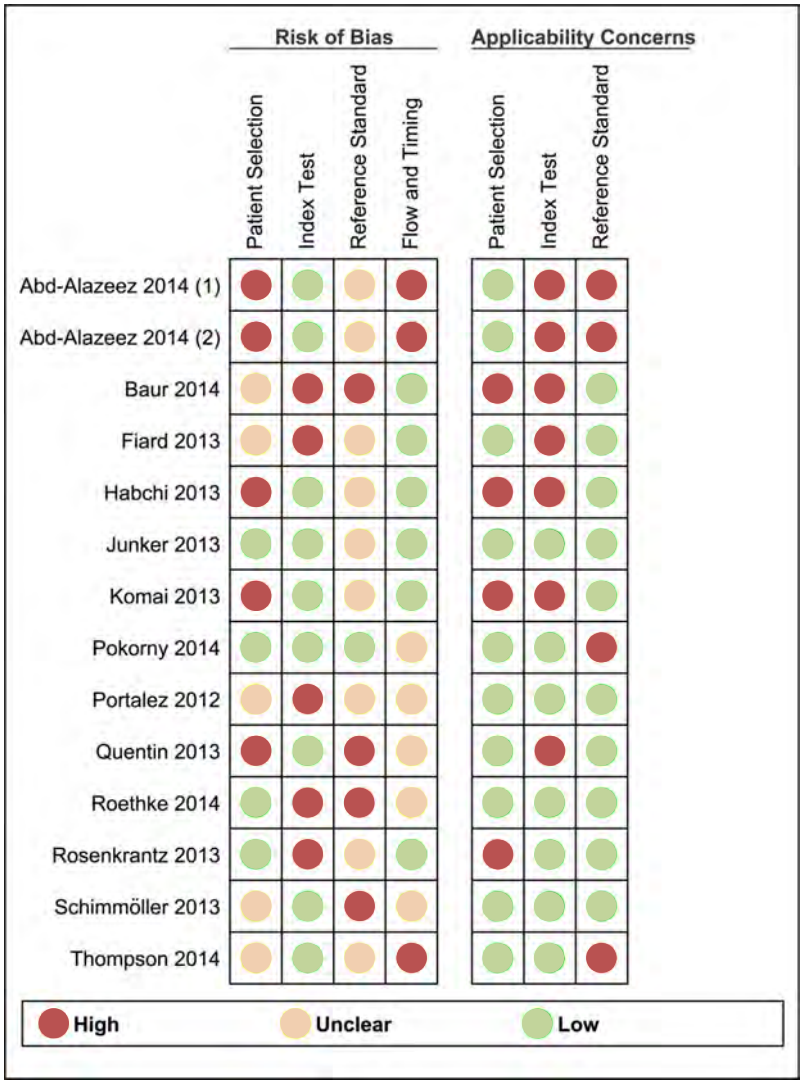


Figure 3 Summary of the risk of bias and applicability concerns. Judgments of the review authors for each domain for each study included as listed in Table 1.

Diagnostic accuracy of the PI-RADS

The pooled sensitivity and specificity of all studies combined was 0.78 (95% CI 0.70–0.84) and 0.79 (95% CI 0.68–0.86), respectively. Figure 4 shows the hierarchic summary ROC plot with 95% CI area and summary point. The negative predictive values ranged from 0.58 to 0.96, and positive predictive values from 0.50 to 0.83.

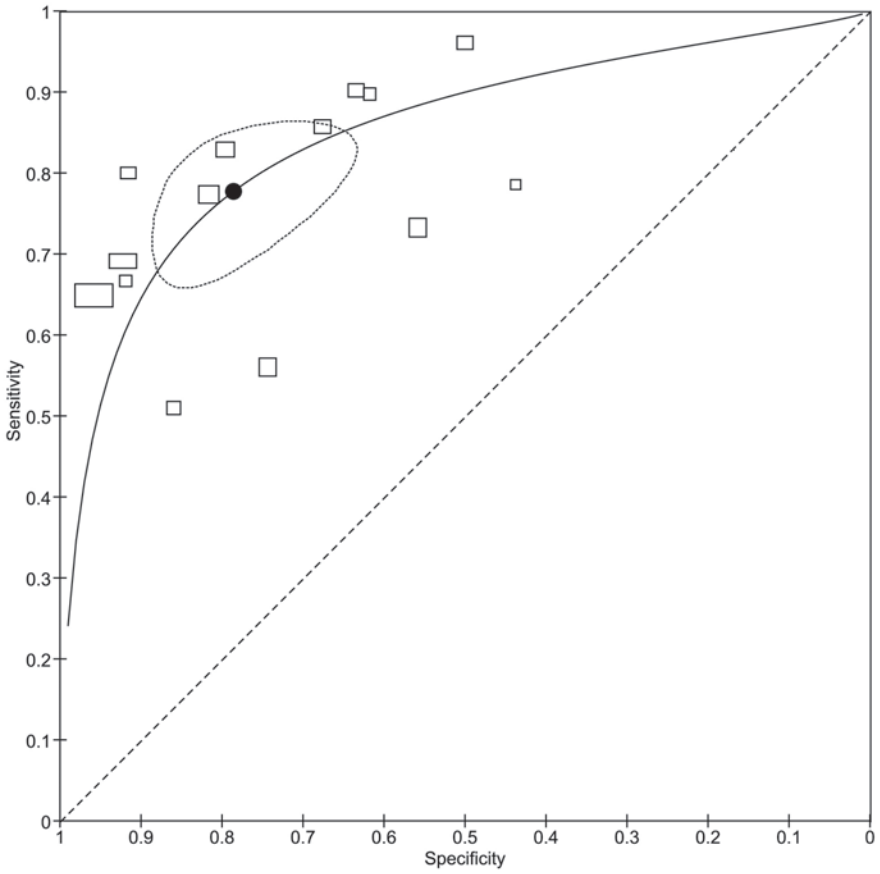


Figure 4 Hierarchic summary receiver operating characteristic (HSROC) (*solid line*) plot with summary point (●) with 95%CI area (*circled area*). Dashed line = no discrimination line (AUC of 0.5, meaning a worthless test), □ =data from individual studies included in meta-analysis (size of square indicates relative size of study population).

Sensitivity analyses

Results of several sensitivity analyses are presented in Figure 5. We were not able to calculate pooled performance values for studies that reported data by region, by lesion, or by zone. Studies with low concerns regarding PI-RADS applicability showed higher sensitivity and specificity (0.82 (95% CI 0.72–0.92) and 0.82 (95% CI 0.67–0.92)) when compared to studies with high concerns for PI-RADS applicability (sensitivity 0.73, 95% CI 0.62–0.82; specificity 0.75, 95% CI 0.61–0.84).

Furthermore, studies with a proper reference standard showed lower sensitivity and specificity (0.76 (95% CI 0.71–0.89) and 0.77 (95% CI 0.64–0.87)) when compared to those

with an inferior reference standard (sensitivity 0.82, 95% CI 0.71–0.89; specificity 0.82, 95% CI 0.64–0.92).

Studies including only patients without previous biopsies had sensitivity of 0.71 (95% CI 0.48–0.86) and specificity of 0.77 (95% CI 0.70–0.83), compared to 0.80 (95% CI 0.65–0.90) and 0.78 (95% CI 0.63–0.89), respectively, in studies including only patients with at least one previous biopsy, and 0.81 (95% CI 0.67–0.90) and 0.72 (95% CI 0.48–0.88), respectively, in studies including a mixed population.

Studies reporting on detection of any PCa as outcome had sensitivity of 0.74 (95% CI 0.67–0.81) and specificity of 0.80 (95% CI 0.70–0.88), compared to 0.84 (95% CI 0.76–0.89) and 0.75 (95% CI 0.66–0.83), respectively, in studies with significant PCaas the outcome.

In studies using an overall 5-point PI-RADS scale, a threshold of 3 had pooled sensitivity of 0.88 (95% CI 0.82–0.93) and specificity of 0.45 (95% CI 0.27–0.65) compared to 0.66 (95% CI 0.54–0.76) and 0.76 (95% CI 0.63–0.85), respectively, in studies using a threshold of 4. In studies using the PI-RADS sum score, sensitivity and specificity of 0.96 (95% CI 0.20–1.0) and 0.73 (95% CI 0.12–0.98) were observed for a threshold of 8, 0.90 (95% CI 0.75–0.96) and 0.66 (95% CI 0.38–0.87) for a threshold of 9, and 0.84 (95% CI 0.74–0.90) and 0.78 (95% CI 0.62–0.89), respectively, for a threshold of 10.

A sensitivity analysis in which we excluded the study of Rosenkrantz et al.,[21] which used radical prostatectomy specimens as the reference standard, showed a sensitivity and specificity that were essentially the same as in the overall analysis (respectively 0.79 and 0.76 with exclusion of this study versus 0.78 and 0.79 without exclusion).

Studies	Pooled performance value (95% CI) Sensitivity	Specificity
All studies	0.78 (0.70 – 0.84)	0.79 (0.68 – 0.86)
Studies accurately using PI-RADS	0.82 (0.72 – 0.89)	0.82 (0.67 – 0.92)
Studies doubtful using PI-RADS	0.73 (0.62 – 0.82)	0.75 (0.61 – 0.84)
Studies that reported data by patient	0.81 (0.72 – 0.87)	0.74 (0.59 – 0.85)
Studies that reported data on any cancer	0.74 (0.67 – 0.81)	0.80 (0.70 – 0.88)
Studies that reported data on significant cancer	0.84 (0.76 – 0.89)	0.75 (0.66 – 0.83)
Studies with a good reference standard	0.76 (0.66 – 0.84)	0.77 (0.64 – 0.87)
Studies with a less good reference standard	0.82 (0.71 – 0.89)	0.82 (0.64 – 0.92)

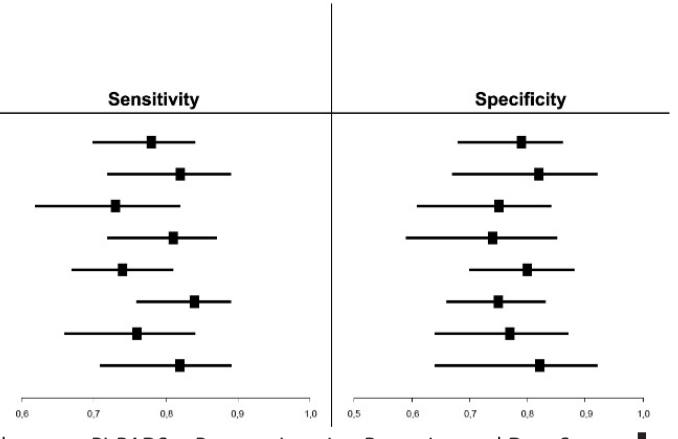


Figure 5 Forest plots of pooled estimates of all studies overall and of different subgroups. PI-RADS = Prostate Imaging Reporting and Data System. ▌ = pooled performance value, horizontal lines are 95% confidence intervals (CIs).

Publication bias

Figure 6 shows the Deeks funnel plot. The statistically nonsignificant value (p=0.36) for the slope coefficient suggests symmetry in the data and a low likelihood of publication bias.

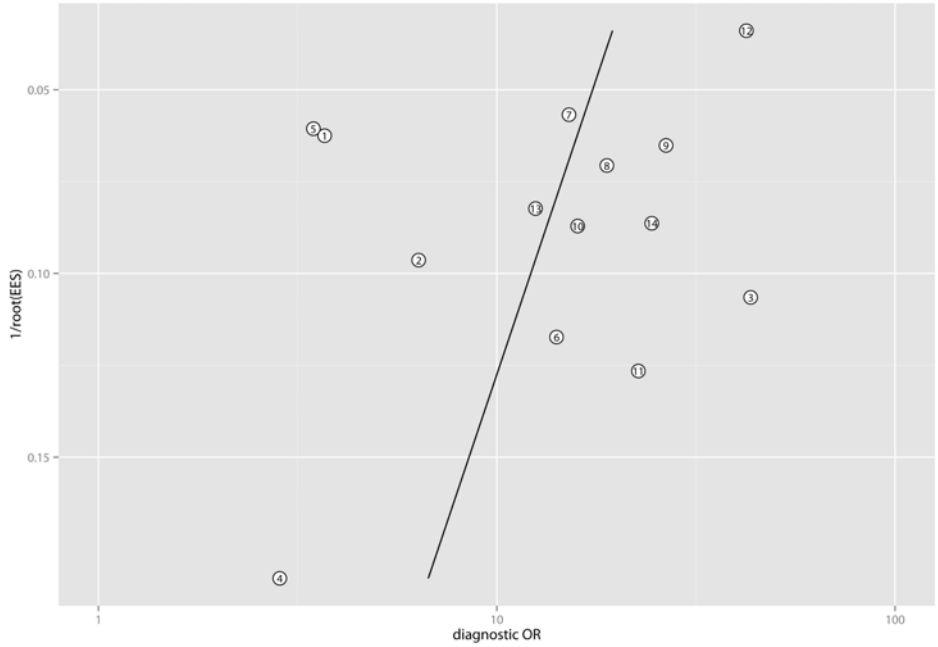


Figure 6 Linear regression test of funnel plot asymmetry. The statistically insignificant value (p= 0.36) for the slope coefficient suggests a low likelihood of publication bias. O = study, —— = regression line. The number in the circle means the study number as listed in Table 1 and 2.

DISCUSSION

This meta-analysis on the diagnostic accuracy of mp-MRI using PI-RADS showed pooled sensitivity and specificity for all studies combined of 0.78 (95% CI 0.70–0.84) and 0.79 (95% CI 0.68–0.86), respectively. The negative predictive values ranged from 0.58 to 0.96, and positive predictive values from 0.50 to 0.83. Although sensitivity of 0.78 (95% CI 0.70–0.84) implies the presence of 22% false-negative patients, the clinical significance of these tumours can be debated. We observed slightly higher sensitivity of 0.84 (95% CI 0.76-0.89) and lower specificity of 0.75 (95% CI 0.66–0.83) for studies with detection of significant PCa as the primary outcome compared to sensitivity of 0.74 (95% CI 0.67–0.81) and specificity of 0.80 (95% CI 0.70–0.88) for studies with detection of all PCa as outcome measure. This

is comparable to the trend observed in four studies in our meta-analysis that performed both analyses[12,13,15,17,20] and can be explained by the presence of more false positive cases and fewer false negative cases.

The studies included showed fairly large heterogeneity regarding calculation of an overall PI-RADS score. Studies with low concerns regarding PI-RADS applicability showed higher sensitivity and specificity (0.82 (95% CI 0.72–0.92) and 0.82 (95% CI 0.67–0.92)) compared to studies with high concerns for PI-RADS applicability (0.73 (95% CI 0.62–0.82) and 0.75 (95% CI 0.61–0.84)). This difference suggests improved accuracy for PCa detection if PI-RADS is used accurately.

So far, only one meta-analysis has been published regarding the diagnostic accuracy of mp-MRI combining T2WI, DWI, and DCE-MRI.[26] That meta-analysis, which included seven studies, showed pooled sensitivity of 0.74 (95% CI 0.66–0.81) and specificity of 0.88 (95% CI 0.82–0.92) for PCa detection. However, only seven studies with 527 patients were included and all studies used other reporting scales. Moreover, studies included in this meta-analysis were all published before publication of PI-RADS. Our results show slightly higher sensitivity and lower specificity compared to that meta-analysis. Three other meta-analyses focused on single modalities and are therefore difficult to compare with our study.[27-29]

The major strength of this study is that it is the first meta-analysis of studies on the currently available mp-MRI PI-RADS. Besides giving an overview of the literature, this will lead to realization among urologists regarding the requirements for better standardization of prostate MRI. Some methodologic issues need to be considered. First, some of the individual studies had limited quality, particularly the studies with a small sample size and inadequate information. The low quality of the studies included may have influenced our meta-analysis outcomes. However, the sensitivity analyses were quite robust. Furthermore, since this is the first meta-analysis of PI-RADS with mp-MRI it provides a new overview of all available evidence at present.

Second, there were high concerns regarding PI-RADS applicability for some studies. We tried to correct for this by performing a sensitivity analysis. This analysis showed a higher accuracy for studies with low concerns compared to those with high concerns regarding PI-RADS applicability. However, it remains unclear whether this difference is fully attributed to PI-RADS applicability.

Third, we found substantial heterogeneity across the PI-RADS studies, with differences in the use of PI-RADS and cut-off values, and types of analyses. This can be explained by a lack of instructions in the ESUR guidelines and by differences in routine clinical practice between different institutions.

Fourth, only one study used whole-gland pathology as the reference, which is considered the gold standard. Although most studies used a combination of targeted and systematic biopsies, the problem remains overcoming the intrinsic limitation of negative biopsies, as these do not exclude the risk of PCa. Furthermore, in case of positive mp-MRI with negative biopsies, no follow-up data were available to evaluate possibly missed biopsies. Diagnostic accuracy might therefore be slightly overestimated.

Fifth, the analyses performed in the individual studies differed; for example, some studies performed a per-patient analysis, whereas others performed a per-lesion or per-core analysis. In studies with a per-patient analysis, no findings on single core positivity or negativity were reported. This makes interpretation of mp-MRI findings difficult, because these studies did not correlate the location of biopsied lesions on mp-MRI with positivity for specific cores. Unfortunately, subgroup analysis using these different types of analyses was not possible because of low numbers.

Sixth, the studies also used different definitions of significant PCa. As in most reviews, we had to pragmatically use the data as presented in the individual studies. Recoding would only be possible using the individual raw data from each study, which were not available. A sensitivity analysis comparing studies that used detection of significant PCa as the outcome measure to studies that used detection of any cancer as the primary outcome showed no substantial difference in pooled estimates of sensitivity and specificity.

Our results show that radiologists increasingly use PI-RADS, as we included one study from 2012, eight from 2013 and five from the first 3 months of 2014. The way in which overall PI-RADS scores are derived and the thresholds applied differ greatly between studies, which might be a consequence of the lack of instruction in the ESUR guidelines. In our results, the sum score appears to perform better than an overall 5-point scale, which might be because of the more objective criteria for calculating a sum score compared to the overall interpretation used in overall 5-point scales. However, both pooled values have an overlapping confidence interval. Owing to the large heterogeneity among studies, we could not provide a recommendation regarding the threshold. However, our

meta-analysis results for mp-MRI with PI-RADS do provide an overview of all available evidence and a better insight into which elements of PI-RADS have to be improved. The ESUR guidelines should be improved; first, clear instructions for calculation of an overall score should be provided. Two frequently used methods in our meta-analysis are the sum score and an overall 5-point scale. A recently proposed method for deriving an overall weighted score from single-modality PI-RADS scores is based on the dominant sequence of the prostate zone for involved lesion.[30] This method needs to be further validated, as it showed promising results in one study.[24] Second, recommendations regarding the use of a threshold for prostate biopsy should be incorporated. Future studies using PI-RADS must be of high quality, with a proper reference standard, a clear description of methodology, and uniformity in the reporting and use of definitions. Moreover, it is important that radiologists are properly trained to use mp-MRI with PI-RADS, and learning curves should be evaluated. Of the studies included, only one contained a statement that radiologists were trained during several weeks.[24] Evaluating inter-reader reproducibility was beyond the scope of our meta-analysis, but reported overall kappa coefficients reported were in the range of 0.340–0.626.[21,23,31] Schimmöller et al. and Quentin et al.[22,32] also evaluated kappa values per modality. Both studies demonstrated a substantially lower kappa for T2WI (0.49-0.55) than for DWI (0.64-0.97) and DCE-MRI (0.65-0.77), which may be because of more subjective interpretation of the criteria of T2WI compared to DWI and DCE-MRI.

The purpose of our study was to review the diagnostic accuracy of mp-MRI with PI-RADS. However, further investigation is required to determine if PI-RADS outperforms other systems for prostate MRI. Once the diagnostic accuracy of PI-RADS is proven to be good, a proper randomized controlled trial should be developed to compare different reporting systems for prostate MRI.

CONCLUSIONS

Pooled data from all included studies that used the PI-RADS with mp-MRI showed sensitivity of 0.78 and specificity of 0.79 for PCa detection. Therefore, it can be concluded that PI-RADS appears to have good diagnostic accuracy in PCa detection. However, no recommendations can be made regarding the use of a threshold because of study heterogeneity.

REFERENCES

1. Siegel R, DeSantis C, Virgo K, et al. Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin* 2012; 62(4): 220-41.
2. Delongchamps NB, Rouanne M, Flam T, et al. Multiparametric magnetic resonance imaging for the detection and localization of prostate cancer: combination of T2-weighted, dynamic contrast-enhanced and diffusion-weighted imaging. *BJU Int* 2011; 107(9): 1411-8.
3. Vilanova JC, Barcelo-Vidal C, Comet J, et al. Usefulness of prebiopsy multifunctional and morphologic MRI combined with free-to-total prostate-specific antigen ratio in the detection of prostate cancer. *AJR Am J Roentgenol* 2011; 196(6): W715-22.
4. Kitajima K, Kaji Y, Fukabori Y, Yoshida K, Suganuma N, Sugimura K. Prostate cancer detection with 3 T MRI: comparison of diffusion-weighted imaging and dynamic contrast-enhanced MRI in combination with T2-weighted imaging. *Journal of magnetic resonance imaging : JMRI* 2010; 31(3): 625-31.
5. Tamada T, Sone T, Higashi H, et al. Prostate cancer detection in patients with total serum prostate-specific antigen levels of 4-10 ng/mL: diagnostic efficacy of diffusion-weighted imaging, dynamic contrast-enhanced MRI, and T2-weighted imaging. *AJR Am J Roentgenol* 2011; 197(3): 664-70.
6. Tanimoto A, Nakashima J, Kohno H, Shinmoto H, Kuribayashi S. Prostate cancer screening: the clinical value of diffusion-weighted imaging and dynamic MR imaging in combination with T2-weighted imaging. *Journal of magnetic resonance imaging : JMRI* 2007; 25(1): 146-52.
7. Iwazawa J, Mitani T, Sassa S, Ohue S. Prostate cancer detection with MRI: is dynamic contrast-enhanced imaging necessary in addition to diffusion-weighted imaging? *Diagnostic and interventional radiology (Ankara, Turkey)* 2011; 17(3): 243-8.
8. Dickinson L, Ahmed HU, Allen C, et al. Scoring systems used for the interpretation and reporting of multiparametric MRI for prostate cancer detection, localization, and characterization: could standardization lead to improved utilization of imaging within the diagnostic pathway? *Journal of magnetic resonance imaging : JMRI* 2013; 37(1): 48-58.
9. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *European radiology* 2012; 22(4): 746-57.
10. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529-36.
11. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005; 58(9): 882-93.
12. Abd-Alazeez M, Ahmed HU, Arya M, et al. The accuracy of multiparametric MRI in men with negative biopsy and elevated PSA level--can it rule out clinically significant prostate cancer? *Urol Oncol* 2014; 32(1): 45 e17-22.
13. Abd-Alazeez M, Kirkham A, Ahmed HU, et al. Performance of multiparametric MRI in men at risk of prostate cancer before the first biopsy: a paired validating cohort study using template prostate mapping biopsies as the reference standard. *Prostate cancer and prostatic diseases* 2014; 17(1): 40-6.
14. Fiard G, Hohn N, Descotes JL, Rambeaud JJ, Troccaz J, Long JA. Targeted MRI-guided prostate biopsies for the detection of prostate cancer: initial clinical experience with real-time 3-dimensional transrectal ultrasound guidance and magnetic resonance/transrectal ultrasound image fusion. *Urology* 2013; 81(6): 1372-8.
15. Habchi H, Bratan F, Paye A, et al. Value of prostate multiparametric magnetic resonance imaging for predicting biopsy results in first or repeat biopsy. *Clinical radiology* 2014; 69(3): e120-8.
16. Junker D, Schafer G, Edlinger M, et al. Evaluation of the PI-RADS scoring system for classifying mpMRI findings in men with suspicion of prostate cancer. *BioMed research international* 2013; 2013: 252939.

17. Komai Y, Numao N, Yoshida S, et al. High diagnostic ability of multiparametric magnetic resonance imaging to detect anterior prostate cancer missed by transrectal 12-core biopsy. *The Journal of urology* 2013; 190(3): 867-73.

18. Portalez D, Mozer P, Cornud F, et al. Validation of the European Society of Urogenital Radiology scoring system for prostate cancer diagnosis on multiparametric magnetic resonance imaging in a cohort of repeat biopsy patients. *European urology* 2012; 62(6): 986-96.

19. Quentin M, Schimmoller L, Arsov C, et al. 3-T in-bore MR-guided prostate biopsy based on a scoring system for target lesions characterization. *Acta radiologica (Stockholm, Sweden : 1987)* 2013; 54(10): 1224-9.

20. Roethke MC, Kuru TH, Schultze S, et al. Evaluation of the ESUR PI-RADS scoring system for multiparametric MRI of the prostate with targeted MR/TRUS fusion-guided biopsy at 3.0 Tesla. *European radiology* 2014; 24(2): 344-52.

21. Rosenkrantz AB, Kim S, Lim RP, 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(2): 482-92.

22. Schimmoller L, Quentin M, Arsov C, et al. Inter-reader agreement of the ESUR score for prostate MRI using in-bore MRI-guided biopsies as the reference standard. *European radiology* 2013; 23(11): 3185-90.

23. 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. *The Journal of urology* 2014.

24. Pokorny MR, de Rooij M, Duncan E, et al. Prospective study of diagnostic accuracy comparing prostate cancer detection by transrectal ultrasound-guided biopsy versus magnetic resonance (MR) imaging with subsequent MR-guided biopsy in men without previous prostate biopsies. *Eur Urol* 2014; 66(1): 22-9.

25. Baur AD, Maxeiner A, Franiel T, et al. Evaluation of the prostate imaging reporting and data system for the detection of prostate cancer by the results of targeted biopsy of the prostate. *Investigative radiology* 2014; 49(6): 411-20.

26. de Rooij M, Hamoen EH, Futterer JJ, Barentsz JO, Rovers MM. Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis. *AJR Am J Roentgenol* 2014; 202(2): 343-51.

27. Zhang ZX, Yang J, Zhang CZ, et al. The value of magnetic resonance imaging in the detection of prostate cancer in patients with previous negative biopsies and elevated prostate-specific antigen levels: a meta-analysis. *Academic radiology* 2014; 21(5): 578-89.

28. Wu LM, Xu JR, Ye YQ, Lu Q, Hu JN. The clinical value of diffusion-weighted imaging in combination with T2-weighted imaging in diagnosing prostate carcinoma: a systematic review and meta-analysis. *AJR Am J Roentgenol* 2012; 199(1): 103-10.

29. Haghighi M, Shah S, Taneja SS, Rosenkrantz AB. Prostate cancer: diffusion-weighted imaging versus dynamic-contrast enhanced imaging for tumor localization-a meta-analysis. *Journal of computer assisted tomography* 2013; 37(6): 980-8.

30. Vache T, Bratan F, Mege-Lechevallier F, Roche S, Rabilloud M, Rouviere O. 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(2): 446-55.

31. 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. *AJR Am J Roentgenol* 2013; 201(4): W612-8.

32. Quentin M, Arsov C, Rohlen S, et al. Inter-reader agreement of multi-parametric MR imaging for the detection of prostate cancer: evaluation of a scoring system. *Rofo* 2012; 184(10): 925-9.

Appendix I Full search strategy

((((((((((((((((prostate[Title/Abstract]) OR prostatic[Title/Abstract]) OR PCa[Title/Abstract]) OR prostate cancer[Title/Abstract]) OR prostatic cancer[Title/Abstract]) OR prostate neoplasm[Title/Abstract]) OR prostatic neoplasm[Title/Abstract]) OR prostate tumor[Title/Abstract]) OR prostatic tumor[Title/Abstract]) OR prostate carcinoma[Title/Abstract]) OR prostatic carcinoma[Title/Abstract])) OR "Prostatic Neoplasms"[Mesh])) AND ((((((((ESUR[Title/Abstract]) OR urogenital radiology[Title/Abstract]) OR PIRADS[Title/Abstract]) OR PI-RADS[Title/Abstract]) OR (reporting[Title/Abstract])) OR likert scales[Title/Abstract]) OR scoring system[Title/Abstract]) OR score[Title/Abstract])))) AND (((((((MR[Title/Abstract]) OR NMR[Title/Abstract]) OR MRI[Title/Abstract]) OR NMRI[Title/Abstract]) OR magnetic resonance imaging[Title/Abstract])) OR "Magnetic Resonance Imaging"[Mesh])

Supplemental Table 1 PI-RADS scoring system (version 1, 2012).

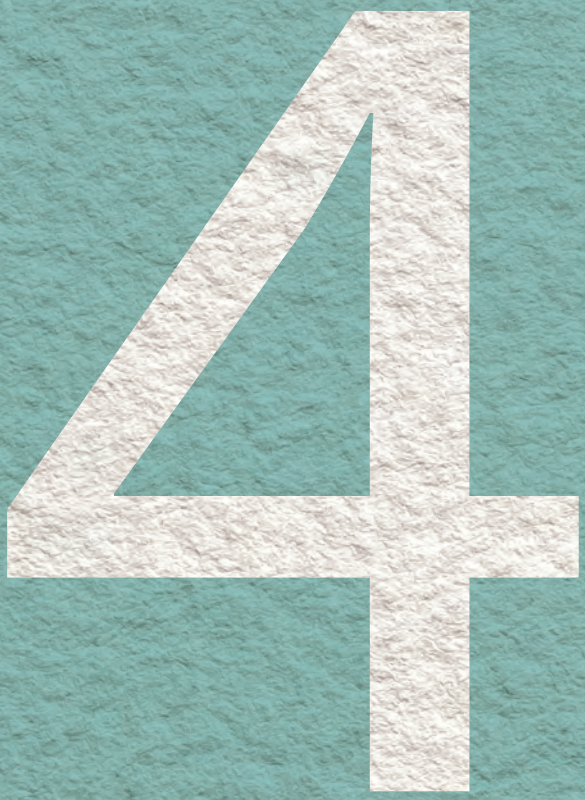
	T2WI peripheral zone (PZ)	T2WI transition zone (TZ)	DWI	DCE-MRI	Overall score
1	Uniform high signal intensity (SI)	Heterogeneous TZ adenoma with well-defined margins	No reduction in ADC. No increase in SI on any high b-value image (≥b800)	Type 1 enhancement curve	Significant PCa highly unlikely to be present
2	Linear, wedge shaped, or geographic areas of lower SI, usually not well demarcated	Areas of more homogeneous low SI, well emarginated, originating from the TZ/BPH	Diffuse, hyper SI on ≥b800 image with low ADC. No focal features. Linear, triangular or geographical features allowed	Type 2 enhancement curve	Significant PCa unlikely to be present
3	Intermediate appearances	Intermediate appearances	Intermediate appearances	Type 3 enhancement curve	Presence of significant PCa equivocal
4	Discrete, homogeneous low signal focus/mass confined to the prostate	Areas of more homogeneous low SI, ill defined	Focal area(s) of reduced ADC but iso-intense SI on high b-value images (≥b800)	+1 For focal enhancing lesion with curve type 2–3	Significant PCa likely to be present
5	Discrete, homogeneous low SI focus with extra-capsular extension/invasive behaviour or mass effect on the capsule (bulging), or broad (>1.5 cm) contact with the surface	Same as 4, but involving the anterior fibromuscular stroma or the anterior horn of the PZ, usually lenticular or water-drop shaped	Focal area/mass of hyper SI on the high b-value images (≥b800) with reduced ADC	+1 For asymmetric lesion or lesion at an unusual place with curve type 2–3	Significant PCa highly likely to be present

Abbreviations: ADC = apparent diffusion coefficient; BPH = benign prostatic hyperplasia; DCE-MRI = dynamic contrast-enhanced magnetic resonance imaging; DWI = diffusion weighted imaging; PCa = prostate cancer; PZ = peripheral zone; SI = signal intensity; T2WI = T2-weighted imaging; TZ = transition zone.

Supplemental Table 2 Technical characteristics.

Study (year)	MRI model (manufacturer)	Field strength (Tesla)	Coil	Endorectal coil
Abd-Alazeez (2013)[13]	Avanto (Siemens) Achieva (Philips)	1.5 3	Multichannel pelvic phased-array coil	No
Abd-Alazeez (2014)[12]	Avanto (Siemens) Achieva (Philips)	1.5 3	Multichannel pelvic phased-array coil	No
Baur (2014)[25]	Magnetom Avanto (Siemens)	1.5	4-channel pelvic phased-array coil	Yes
Fiard (2013)[14]	Achieva (Philips)	3	32-channel phased-array coil	No
Habchi (2013)[15]	Discovery MR750 (General Electric)	3	Pelvic phased-array coil	No
Junker (2013)[16]	Magnetom Skyra (Siemens)	3	18-channel phased-array body coil	No
Komai (2013)[17]	Intera Achieva (Philips)	1.5	4-channel body coil	No
Pokorny (2014)[24]	Magnetom Skyra (Siemens)	3	NR	No
Portalez (2012)[18]	Achieva (Philips) Avanto (Siemens)	1.5 1.5	Pelvic phased-array coil and SENSE cardiac coil 8-channel pelvic phased-array coil	Yes, in 1 of 2 centres No
Quentin (2013)[19]	Magnetom Trio (Siemens)	3	6-channel phased-array body coil	No
Roethke (2014)[20]	Magnetom Trio (Siemens)	3	Multi-channel body coil and integrated spine phased-array coil	No
Rosenkrantz (2013)[21]	Magnetom Trio (Siemens)	3	Pelvic phased-array coil	No
Schimmoeller (2013)[22]	Magnetom Trio (Siemens)	3	6-channel phased-array body coil	No
Thompson (2014)[23]	NR	1.5 and 3	14-channel spine and 18-channel pelvic phased array coils	No

Abbreviations: MRI = magnetic resonance imaging; NR = not reported.



Accuracy of MRI for local staging of prostate cancer: a diagnostic meta-analysis

Published as

de Rooij M, Hamoen EHJ, Witjes JA, Barentsz JO, Rovers MM

Accuracy of magnetic resonance imaging for local staging of prostate cancer: a diagnostic meta-analysis

European Urology, 2016 Aug;70(2):233-45

ABSTRACT

Context

Correct assessment of the tumour stage is crucial for prostate cancer (PCa) management.

Objective

To assess the diagnostic accuracy of magnetic resonance imaging (MRI) for local PCa staging and explore the influence of different imaging protocols.

Evidence acquisition

We searched PubMed, Embase and Cochrane from 2000 up to August 2014. We included studies that used MRI for detection of extracapsular extension (ECE; T3a), seminal vesicle invasion (SVI; T3b), or overall stage T3 PCa, with prostatectomy as the reference standard. Methodologic quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies tool by two independent reviewers. Data necessary to complete 2x2 tables were obtained, and patient, study, and imaging characteristics were extracted. Accuracy was reported for the most experienced or first reader. Results were pooled and plotted in summary receiver operating characteristics plots.

Evidence synthesis

A total of 75 studies (9,796 patients) could be analysed. Pooled data of ECE (45 studies, 5,681 patients), SVI (34 studies, 5,677 patients), and overall stage T3 detection (38 studies, 4,001 patients) showed sensitivity and specificity of 0.57 (95% confidence interval (CI) 0.49-0.64) and 0.91 (95% CI 0.88-0.93), 0.58 (95% CI 0.47-0.68) and 0.96 (95% CI 0.95-0.97), and 0.61 (95% CI 0.54-0.67) and 0.88 (95% CI 0.85-0.91), respectively. Functional imaging in addition to T2-weighted imaging, and use of higher field strengths (3 Tesla) improved sensitivity for ECE and SVI. Sensitivity of ECE was not improved by endorectal coil use.

Conclusions

MRI has high specificity but poor and heterogeneous sensitivity for local PCa staging. An endorectal coil showed no additional benefit for ECE detection, but slightly improved sensitivity for SVI detection. Higher field strengths and the use of functional imaging techniques can slightly improve sensitivity.

Patient summary

We pooled the results of all previous studies that evaluated MRI for detection of tumour growth outside the prostate. MRI is not sensitive enough to find all tumours with extraprostatic growth.

INTRODUCTION

Prostate cancer (PCa) has emerged as the most common malignancy among Western males, and is the second leading cause of cancer-related mortality.[1] Traditionally, PCa detection and local staging depend on a combination of diagnostic tests. Serum prostate specific antigen (PSA) and digital rectal examination (DRE) are used to identify men who need subsequent transrectal ultrasound guided biopsies (TRUS-GB).[2] Although these techniques are able to detect PCa and estimate disease aggressiveness, they often underestimate tumour stage and are not accurate for detection of locally advanced disease.[3,4] Correct assessment of the tumour stage is crucial for disease management. Curative treatment is most likely when the TNM stage is ≤T2c, that is, when extracapsular extension (ECE, stage T3a), seminal vesicle invasion (SVI, stage T3b), and metastatic disease (N+ and/or M+) are not present. Magnetic resonance imaging (MRI) is increasingly used to aid prostate biopsy targeting and, more accurate detection of PCa.[5,6] MRI can also improve the determination of the tumour extent.[4] Many studies have investigated the accuracy of MRI in local staging. Studies differ in their use of magnetic field strengths, the use of an endorectal coil, and combinations of anatomic and functional MRI techniques. Heterogeneous results are driving the ongoing debate regarding the usefulness of MRI and the best imaging protocol for PCa staging.

So far, two meta-analyses on local staging accuracy have been published. Engelbrecht et al. included literature up to 2000, and a more recent survey by Silva et al. restricted to studies using 1.5 Tesla (T) devices with an endorectal coil.[7] Since 2000, local staging of PCa has been studied extensively using many different imaging protocols. An updated search without limitations on field strength or coil use is warranted to provide a comprehensive overview of evidence available for the most common imaging protocols. The aim of this meta-analysis was therefore to assess the diagnostic accuracy of MRI for local staging and to analyse the influence of different imaging protocols in men with biopsy-proven PCa, with radical prostatectomy specimens as the reference standard.

EVIDENCE ACQUISITION

We performed this systematic review and meta-analysis according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis statement.[8]

Literature search

We performed a systematic search in PubMed, the Cochrane Central Register of Controlled Trials, and Embase for studies evaluating the diagnostic accuracy of MRI for local staging of PCa. Our search strings combined 'prostate cancer' synonyms with synonyms for 'MRI' and 'staging'. Reference lists of the included articles and reviews were checked, and related articles were traced to complement the electronic query. Searches were performed from 2000 up to August 12, 2014, and were restricted to publications in English. The bibliographic database of EndNote X5 (Thomas Reuters, New York City, NY) was used to filter duplicate articles.

Study selection

We included studies if (1) accuracy was assessed for local staging (ECE/T3a, SVI/T3b, or overall stage T3 disease when there was no stratification between T3a and T3b) using MRI as the index test in patients with biopsy-proven PCa; (2) radical prostatectomy was used as the reference standard; and (3) we could reconstruct two-by-two tables of ECE, SVI, and/or overall stage T3. We excluded studies that focused on restaging or lymph node or bone staging, or used other imaging techniques for local staging, e.g. PET/CT, TRUS-GB, and CT. Two reviewers (MdR/EH) independently assessed the eligibility of the identified papers. Any disagreements were resolved by discussion with a third reviewer (MMR).

Data extraction

We extracted data on patient, study, and imaging characteristics for all included studies. Patient characteristics comprised age, PSA level, Gleason score for biopsy tissue, clinical risk group, and prevalence of ECE, SVI, or overall stage T3 disease. Study characteristics included design, sample size, inclusion and exclusion criteria, ECE/SVI criteria, number of readers, reader experience, consensus reading, radical prostatectomy technique, reference standard technique, consecutive patient selection, blinding for reference and index tests, interval between biopsy and MRI, and interval between MRI and radical prostatectomy. Imaging characteristics included the device manufacturer, device model, magnetic field strength in Tesla (T), type of coil, and imaging sequence details, that is, T2-weighted imaging (T2WI), dynamic contrast-enhanced MRI (DCE-MRI), diffusion weighted imaging (DWI), and/or magnetic resonance spectroscopic imaging (MRSI). One reviewer extracted data for all included studies using a standardized data extraction form. A second reviewer was contacted to resolve unclear issues by consensus.

Methodological quality assessment

We assessed the risk of bias and the applicability at study level using the validated Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) scoring system.[9] Four domains are scored: (1) patient selection, which describes the method for patient selection and the patients included; (2) index test, which describes the test being studied and how it was conducted and interpreted; (3) reference standard, which describes the reference standard test used and how it was conducted and interpreted; and (4) flow and timing, which describe the flow of patient inclusion and exclusion and the interval between the index test and the reference standard. The quality assessment was performed by two independent reviewers (MdR and EH). Any disagreements were resolved by discussion with a third reviewer (MMR).

Data synthesis and analysis

Data from each study were summarized in 2 x 2 tables of true positive, false positive, true negative, and false negative values to calculate sensitivity and specificity values for ECE, SVI, and/or overall stage T3 detection. If analyses of different imaging protocols were performed within a single study, for instance T2WI and T2WI + DWI + DCE-MRI, we chose the most up-to-date technique (T2WI + DWI + DCE-MRI). In the case of different cut-off thresholds, we used the clinically most appropriate one, and when accuracy was reported for different readers we chose the most experienced reader, or the first reader when experience was not described. If staging accuracy had been assessed on a patient basis (eg, ECE/SVI/T3 present) and a region basis (eg, hemi-prostates, sextants) within a study, we included the patient-based analysis to approximate the clinical practice. Authors of studies that did not report sufficient data were asked to provide additional information. To graphically display the sensitivity and specificity at study level, we used Review Manager 5 software from the Cochrane collaboration. We drew forest plots to show variation and to explore heterogeneity for sensitivity and specificity, and plotted the results on a receiver operating characteristic (ROC) curve. Primary outcomes are pooled estimates of sensitivity and specificity with 95% confidence intervals (CIs). We used the Metadas tool within the statistical software package SAS to carry out the meta-analyses.[10] The analyses were imported into RevMan5 (The Nordic Cochrane Center, Copenhagen, Denmark) and used for fitting hierarchic summary receiver operating characteristic (HSROC) plots. Publication bias was studied for ECE, SVI, and overall stage T3 separately using Deeks funnel plots. This analysis was performed using the R statistical package system (R Foundation for Statistical Computing, Vienna, Austria). To study the heterogeneity, we performed sensitivity analyses of several clinically relevant covariates: analysis method (patient or region level), use of an endorectal coil, field strength of magnet (1.0, 1.5, or 3.0T), use of functional imaging techniques in addition to T2WI (DCE-MRI, DWI, and/or MRSI), number of participants (less

or more than 50), absolute prevalence of ECE/SVI/overall stage T3 (less or more than 10%), QUADAS applicability risk (high risk absent or present), and risk category for the study population (low/mixed/high/unclear risk).

EVIDENCE SYNTHESIS

Literature search

Figure 1 provides an overview of the literature search and study selection. Our search yielded 4682 unique records, of which 315 remained after screening titles and abstracts. The full-text of these studies was reviewed for eligibility. Studies were excluded if no staging accuracy was reported, when we were unable to reconstruct a 2x2 contingency table, or when the study was not in English. Manual checking of references cited in the included studies and relevant review articles resulted in two additional papers. This yielded 75 studies for inclusion, of which 45 studies reported on ECE (5,681 patients), 34 on SVI (5,677 patients), and 38 studies that did not stratify between ECE and SVI, but only reported on overall stage T3 detection (4,001 patients).[11-85]

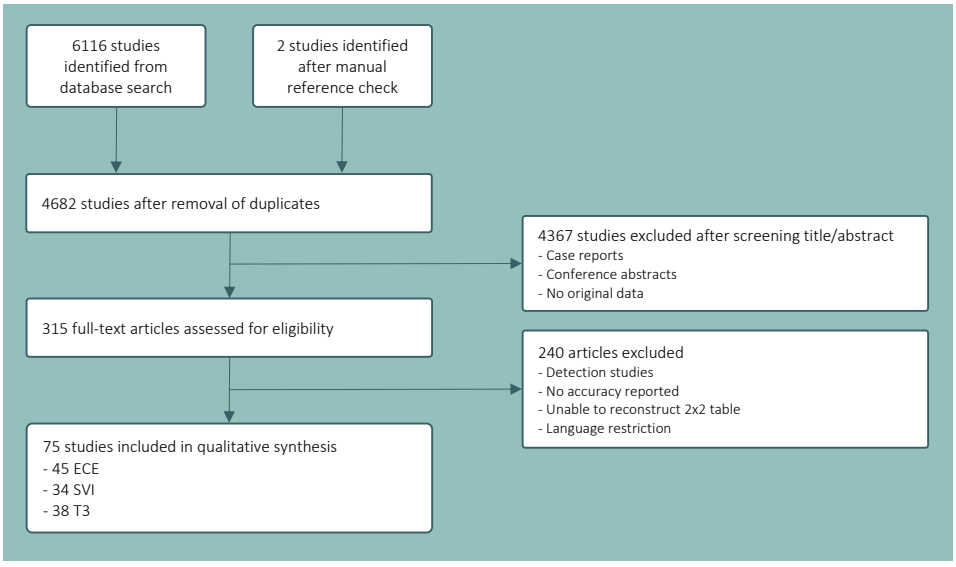


Figure 1 Flow chart of the study inclusion.

Study characteristics

Table 1 summarizes the patient, technical, and study characteristics: patient age, PSA level, Gleason score, prevalence of ECE/SVI/T3 disease, imaging protocol details, and study design details. Table 2 provides the characteristics of the included studies separately.

Table 1 Summary of patient-, study-, and imaging characteristics.

Patient characteristics (range)		Study characteristics (no.)	
Median age in years	58 – 68	Study design	
Median PSA level in ng/mL	4.9 – 14.1	Cohort	72
Biopsy Gleason Score	3 – 10	Case-control	3
Prevalence of ECE (%)	9 – 81	Image interpretation	
Prevalence of SVI (%)	1 – 27	Prospective	18
Prevalence of T3 disease (%)	15 – 81	Retrospective	42
		Combination	5
Technical characteristics (no.)		Not reported	10
Field strength (in Tesla)		Criteria ECE/SVI/T2T3	
1.0	2	Description	45
1.5	47	Description & likelihood scale	18
3.0	21	Reference to guideline only	2
Both 1.5 & 3.0	4	Not reported	10
Not reported	1	Radiologist blinded to clinical data	
Endorectal coil		Yes	35
Yes	47	No	14
No	24	Part of the readers blinded	1
Part	3	Not reported	25
Not reported	1	Pathologist blinded to MRI findings	
Functional techniques		Yes	18
T2WI only	41	No	2
1 additional functional technique	21	Not reported	55
2 additional functional techniques	12		
3 additional functional techniques	1		

Abbreviations: ECE = extracapsular extension; MRI = magnetic resonance imaging; PSA = prostate specific antigen; SVI = seminal vesicle invasion; T2T3 = T3 disease; T2WI = T2 weighted imaging.

Table 2 Patient-, study-, and imaging characteristics of all included studies separately.

First author (year)	Patient characteristics						Study characteristics								Imaging characteristics		
	Age (y)		PSA (ng/ml)		Gleason Score		n=	Design	Reading	Outcome	Analysis	Readers	Criteria	Blinding	Tesla	mp-MRI	erc
	Mean	Range	Mean	Range	Mean	Range											
Akin (2003)	58	45-71	NR	NR	NR	NR	25	cohort	P	E, S, T3	patient	2	description	y	1.0	-	y
Akin (2006)	60*	41-75	5.9	2.1-28	Cat	6-8	148	cohort	R	E	patient[d]	2	no description	y	1.5	-	y
Allen (2004)	NR	NR	NR	NR	NR	NR	55	cohort	both	T3	patient	4	description	y	1.5	-	n
Armitage (2013)	64.1	46-74	NR	NR	NR	NR	35	cohort	NR	T3	patient	NR	no description	NR	1.5	-	n
Augustin (2009)	62.8	47-71	8.9	1.7-31.7	NR	NR	27	cohort	P	E	patient	1	description	n	3	-	n
Bernstein (2000)	60.2	44-71	8.3	1-33.6	Cat	NR	124	cohort	R	E	patient	NR	NR	NR	1.5	-	y
Beyersdorff (2005)	62	50-72	7.5	2-14	NR	NR	22	cohort	P	T3	patient	2¶	description	y	both	-	both
Bloch (2007)	65	42-78	9.3	1-42.8	6	3-9	32	cohort	P	E	patient	2	description	y	1.5	DCE	y
Bloch (2012)	58.5	47-72	10.6	2-117	7*	6-9	108	cohort	both	E	patient	6	description	n	3	DCE	y
Borre (2005)	NR	NR	NR	NR	NR	NR	38	cohort	NR	E	patient	NR	description	NR	1.5	-	n
Brajtbord (2011)	59.3	NR	6.6	NR	NR	6-10	179	cohort	R	T3	patient	>2	no description	NR	NR	-	y
Brassell (2004)	62	NR	14.8	NR	NR	NR	40	cohort	P	T3	patient	1	description	y	1.5	-	y
Brown (2009)	58	42-73	6.1*	0.5-30.5	Cat	6-10	62	cohort	P	T3	patient	team	no description	n	1.5	-	y
Cerantola (2013)	67	NR	12.7	NR	Cat	NR	60	cohort	R	T3	patient	2	description	n	3	DWI DCE	y
Chandra (2007)	60*	44-72	6.3*	2-82	6*	5-8	38	cohort	R	E, S, T3	patient	2	5-point scale	n	1.5	-	y
Chong (2014)	67.7	47-76[y]	15.3	0.1-57.0[y]	7*	6-9	76[a]	c-c	R	E	region	2¶	5-point scale	y	3	MRSI	n
Cornud (2000)	65	NR	14.9	5-30	6	NR	94	cohort	NR	T3	patient	2¶	description	n	1.5	-	y
Cornud (2002)	NR	NR	14.5	3-66	6	NR	336	cohort	R	E, S, T3	patient	2¶	description	n	1.5	-	y
Cornud (2012)	63*	47-76	7*	2.8-23	NR	NR	178	cohort	P	E	patient	2¶	3 point scale	y	1.5	-	y
Counago (2014)	61	NR	Cat	NR	Cat	NR	47	cohort	R	T3	patient	1	description	NR	3	DWI DCE	n
Fütterer (2005)	63*	42-74	7.8*	3.7-78	6*	3-9	103	cohort	P	E, S	patient	3	5-point scale	y	1.5	DCE	y
Fütterer (2006)	65.4*	41-78	14.1*	4 to 78	6*	3-9	32	cohort	P	T3	both	3	5-point scale	y	3	-	y
Fütterer (2007)	62.2*	51-72	8.9*	1 to 45	6*	4-7	81[b]	cohort	P	E, S, T3	patient	5	5-point scale	n	1.5	-	y
Graser (2007)	63	NR	11.5	3.1 to 73	6*	4-9	106	cohort	R	E	both	3	NR	y	1.5	-	y
Gupta (2014)	60.1*	43-74	6.9	1.2-46.3	Cat	NR	60	cohort	R	E, T3	patient	1	NR	y	3	DWI DCE MRSI	y
Hara (2013)	NR	NR	Cat	NR	Cat	NR	132	cohort	R	E	region	2	description	NR	3	-	y
Hegde (2013)	57.8*	56-65[y]	5.8	3.1-10.3	Cat	6-10	118	cohort	R	E, S, T3	patient	team	description	n	3	DWI DCE MRSI	y
Heijmink (2007)	61	51-70	7.8	3.5-24.6	NR	5-9	46	cohort	P	E, S, T3	patient	4	description	y	3	-	y
Hole (2013)	62	43-78	19.8	1.4-288	Cat	5-10	208	cohort	P	T3	patient	2	description	NR	1.5	MRSI	n
Hwii Ko (2011)	62.8	46-74	9.6	0.4-24.4	7	4-10	121	cohort	NR	E, S	patient	1	description	y	3	-	n
Ikonen (2001)	62	55-74	12	0.1-51	NR	NR	44	cohort	R	E, S	region	2	description	y	1.5	-	y
Jeong (2013)	66.1	42-85	17.8	0.3737	Cat	NR	922	cohort	P	E, S	patient	4	description	n	both	DWI	both
Johnston (2013)	62	35-74	8.7	NR	Cat	NR	568	cohort	R	E, S	patient	team¶	description	NR	1.5	-	n
Jung (2008)	64.5	42-77	7.7*	2.4-67	NR	6-10	217	cohort	R	S	patient	2	5-point scale	y	1.5	-	y
Kim (2008)	66*	44-82	7.6*	1.5-138	7*	5-10	166[c]	c-c	R	S	patient	2	5-point scale	y	3	MRSI	n
Kim (2010)	59.1	NR	Cat	NR	Cat	NR	32	cohort	R	E, S	region	1	description	y	1.5	-	n
Kim (2012)	64.8	47-76	11.7	3.0-37.0	NR	6-9	63	cohort	R	E, S	patient	2¶	description	NR	3	-	both
Kwek (2004)	62.5	50-74	15*	5.5-65.9	7*	5-8	21	cohort	NR	T3	patient	1	description	NR	1.5	-	n
Latchamsetty (2007)	59	43-75	6.9*	0.5-115	NR	NR	40	cohort	P	E, S	patient	4	description	n	1.5	-	y

Lee (2010)	63	44-76	NR	2.9-33.5	7	4-9	91	cohort	R	E, S	patient	2¶	description	NR	1.5	MRSI	both
May (2001)	66	54-73	11.8	NR	NR	NR	54	cohort	P	E, S, T3	patient	2	description	y	1.5	DCE	y
Min (2012)	67.8	NR	13.4	NR	Cat	NR	126	cohort	R	T3	patient	1	description	NR	3	DWI DCE	n
Nakashima (2004)	65.9	52-74	14.2	1.6-153.1	NR	NR	95	cohort	NR	E, S, T3	patient	2	description	n	1.5	DCE	y
Nepple (2011)	61*	48-72	7*	1.3-35	NR	6-10	91	cohort	R	E, S	patient	NR	no description	n	1.5	-	y
Ogura (2001)	68.5*	54-75	7.4*	NR	NR	NR	38	cohort	P	E, S, T3	patient	2	description	y	1.5	DCE	y
Oon (2014)	64	NR	9.1	NR	NR	NR	88	cohort	R	E, S	patient	5	no description	NR	1.5	DWI DCE	n
Otto (2014)	65	53-75	13.5*	3.756	7	6-9	37	cohort	R	E, S	both	2	5-point scale	y	3	DWI DCE MRSI	y
Park (2007)	63	43-74	8.6	2.0-35	7*	5-9	108	cohort	R	E, S, T3	patient	2¶	5-point scale	y	both	-	both
Park (2010)	63.5*	50-72	7.4*	3.5-41.2	NR	6-9	54	cohort	R	S, T3	patient	2¶	description	NR	1.5	DCE	y
Park (2014)	64*	43-73	5.3*	1.7-58.5	Cat	6-10	353	cohort	R	T3	patient	2	5-point scale	y	3	DWI DCE	n
Porcaro (2013)	66	51-77	11	1-87.6	6	4-9	154	cohort	R	E, S, T3	patient	2	description	y	1.5	-	y
Pugh (2012)	60*	42-76	4.9*	0.4-9.9	7	NR	171	cohort	P	T3	patient	1	description	NR	1.5	-	y
Ren (2009)	68	56-84	12.9	4.5-67.0	6	3-10	283	cohort	R	S	patient	2	description	NR	3	DWI	n
Renard-Penna (2011)	60*	39-71	8	1.8-30	NR	NR	101	cohort	P	T3	patient	2	description	y	1.5	DCE	n
Roethke (2012)	62.7	42-77	8.9	0.4-52.5	7*	3-9	385	cohort	R	E	patient	2¶	description	n	1.5	-	y
Roethke (2014)	62.8	43-76	8.8*	1.1-52.5	7*	NR	376	cohort	R	S	patient	2	description	NR	1.5	-	y
Rosenkrantz (2013)	61.1	47-81	6.6	2.3-30.9	NR	NR	51	cohort	R	E	region	2	4-point scale	y	3	DWI	n
Rud (2014)	64	42-74	7.8	1.2-54	NR	NR	199	cohort	P	T3	patient	1	description	n	1.5	DWI	n
Ruprecht (2012)	63	49-78	NR	NR	7*	5-8	46	cohort	R	T3	patient	2	description	y	1.5	-	y
Sala (2006)	59*	44-73¥	5.8*	3.3-72¥	Cat	NR	51	c-c	R	E, S	region	2	5-point scale	y	1.5	-	y
Somford (2013)	62.4	NR	10	NR	Cat	Cat	183	cohort	P	T3	patient	2	reference	n	1.5	DWI DCE	y
Soulié (2001)	64.2	49-74	10.9	1.2-39	6	5-9	176	cohort	P	T3	patient	1	description	y	1.0	-	n
Soylu (2013)	68*	43-75	12.1*	1.5-65	7*	6-9	131	cohort	R	S	both	2	5-point scale	y	1.5	DWI DCE	y
Tan (2008)	61	33-71	NR	NR	NR	NR	32	cohort	R	E	patient	1[a]	description	NR	1.5	-	y
Tanaka (2013)	67*	51-74	7*	2.9-27.6	Cat	6-8	67	cohort	P	T3	region	1	description	NR	3	DWI DCE	n
Torricelli (2006)	66	57-75	7.8	2-28	NR	NR	22	cohort	NR	T3	patient	2	description	NR	both	-	both
Torricelli (2008)	66	52-77	9.2	4-39	7	4-8	42	cohort	NR	E	patient	2	description	NR	3	-	n
Tsao (2013)	68.9	50-85	16.9	0.1-107	Cat	NR	94	cohort	NR	E, S	patient	1	description	NR	1.5	-	y
Wang (2004)	57.5	32-74	7.6	0.7-113.4	NR	5-10	344	cohort	R	T3	patient	10	5-point scale	NR	1.5	MRSI	y
Wang (2007)	59	40-86	NR	NR	NR	NR	255	cohort	R	E, S	patient	2	5-point scale	y	1.5	-	y
Wang (2014)	64*	57–68¥	7*	5.212¥	Cat	NR	25	cohort	R	T3	patient	2	NR	NR	3	DWI DCE	NR
Wetter (2006)	66*	48-78	8.5	3.2-36.1	6*	4-9	38	cohort	both	E	patient	2	description	y	1.5	MRSI	y
Xylinas (2011)	63.8	45-75	11.3	2.9-26	7	5-10	70	cohort	P	E, S, T3	patient	1	description	y	1.5	DCE	n
Zhang (2007)	62.2	NR	8.1	NR	6	NR	110	cohort	P	E, S, T3	patient	1	3-point scale	NR	1.5	-	y
Zhang (2009)	58*	40-76	5.3*	1.5-21	NR	6-9	158	cohort	R	E, S	patient	2	5-point scale	y	1.5	MRSI	y

Abbreviations: Cat = categorical; DCE = dynamic contrast-enhanced MRI; DWI = diffusion weighted imaging; E = extracapsular extension; erc = endorectal coil; mp-MRI = multiparametric technique; MRSI = magnetic resonance spectroscopic imaging; n = no; NR = not reported; P = prospective; PSA = prostate specific antigen; R = retrospective; S = seminal vesicle invasion; T3 = T3 disease; y = yes.

Legend: *median; ¥inter quartile range; ¶consensus reading. [a]Chong 2014: case-control study; 31 patients with ECE, 46 patients without ECE. [b]Fütterer 2007: 81 patients that underwent staging MRI with and without endorectal coil. Results reported separately for both groups. [c]Kim 2008: case-control study; 30 patients with SVI, 136 patients without SVI. [d]Akin 2006: transition zone tumours only.

Quality assessment

Overall, the quality was moderate (Figure 2 and Supplemental Figure 1). In particular, the risk of bias was unclear for many studies because of a lack of reporting on patient enrolment, on blinding to the index test during interpretation of the reference test, and on the interval between the index test and surgery. In the patient selection domain, seven studies had a high risk of bias because of a case-control design or inappropriate exclusion criteria. Three studies were assigned high concern regarding applicability to our review question because of unclear patient selection and study design, or because only transition zone tumours were analysed. For the index test domain, three studies had a high risk of bias. Two of these studies used a likelihood scale for assessment of tumour extension, but did not provide a cut-off level. In one study, the readers were not blinded to the reference test when interpreting the index test. Thirteen of the 75 studies showed high concern regarding applicability of the index test because there was no information provided on index test characteristics, readers, blinding, and/or interpretation of the MR images. Other studies used a functional technique in only a subset of patients without reporting the results separately. One study performed both 1.5T and 3.0T imaging in the same patients and used both techniques together during interpretation. For the reference test domain, only one study had a high risk of bias because the reference standard was interpreted without blinding to the index test. All the studies used prostatectomy specimens as the reference standard, so there were no concerns regarding the applicability for this domain. For the flow and timing domain, three studies had a high risk of bias. One study with high risk of bias did not avoid inappropriate exclusions and the staging accuracy was dependent on the accuracy of tumour detection in the first step of the protocol. The two remaining studies did not use the same reference standard for all patients. No studies were excluded from the analysis on the basis of the quality assessment.

Diagnostic accuracy

Figure 3 and Supplemental Figure 2 show hierarchic summary ROC plots with summary point and 95% CI areas of ECE, SVI, and overall stage T3 detection. The pooled sensitivity and specificity for ECE, SVI, and overall stage T3 were 0.57 (95% CI 0.49-0.65) and 0.91 (95% CI 0.88-0.93), 0.58 (95% CI 0.47-0.68) and 0.97 (95% CI 0.95-0.98), and 0.61 (95% CI 0.54-0.67) and 0.88 (95% CI 0.85-0.91), respectively.

Sensitivity analyses

Figure 4 and Supplemental Figure 3 show the results of the sensitivity analyses performed for subgroups of studies to explore the influence of patient characteristics, methodological differences, and technical details on pooled sensitivity and specificity estimates. Overall, specificity estimates showed comparable results, but differences in sensitivity were observed in several sensitivity analyses.

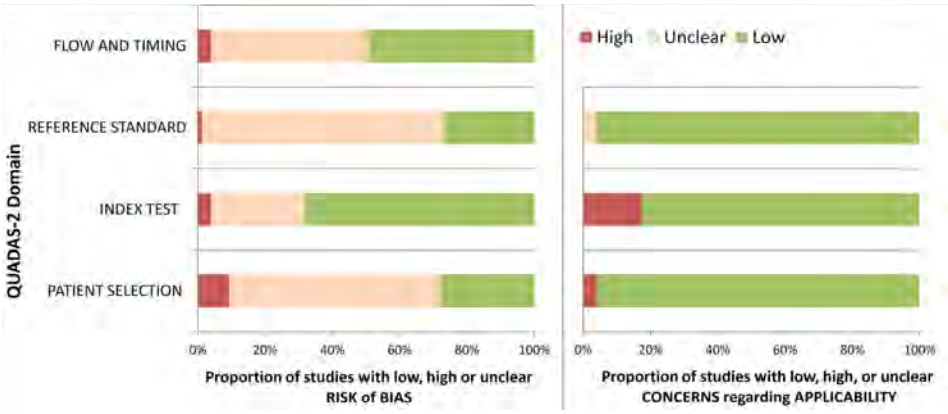


Figure 2 Risk of bias and applicability concerns graph: review of authors' judgements about each domain presented as percentages across included studies.

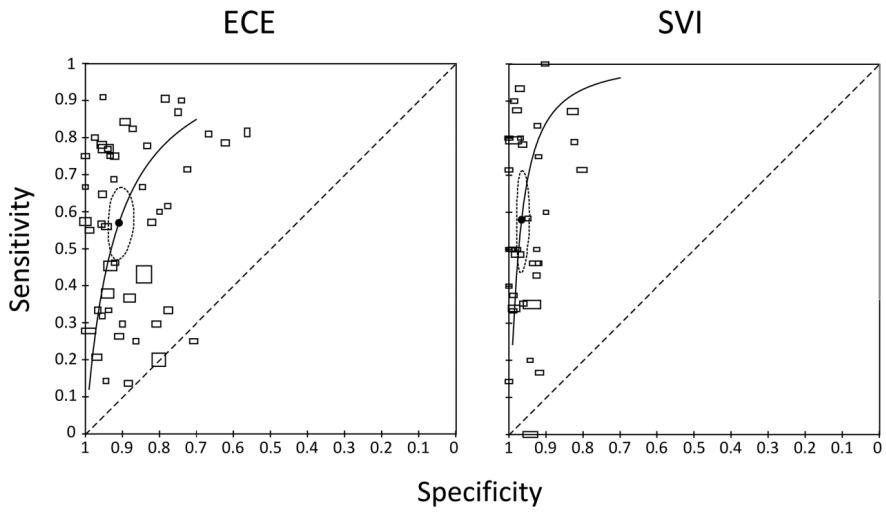


Figure 3 Hierarchic summary receiver operating characteristic (HSROC) plots (solid line) with summary point and 95% confidence interval (CI) areas (circled area). Dashed line = no discrimination line (AUC of 0.5, meaning a worthless test), □ =data from individual studies included in meta-analysis for extracapsular extension (ECE), and seminal vesicle invasion (SVI). The size of the squares corresponds with the relative size of the study population.

Sensitivity analyses for ECE (stage T3a) detection

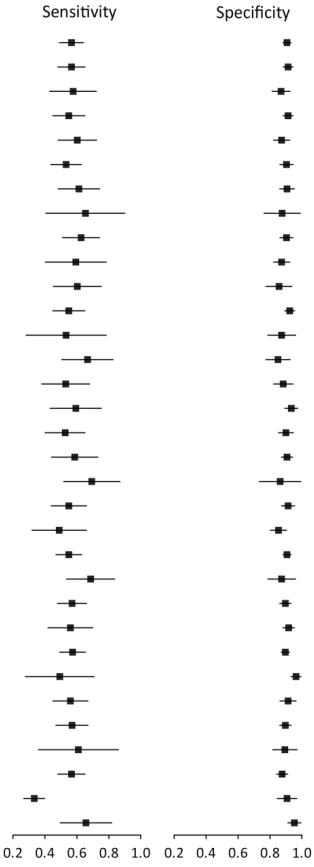
Figure 4a shows the results for ECE (T3a) detection. Studies with patient-level analysis had lower sensitivity (42 studies; 0.55, 95% CI 0.47-0.63) than studies reporting at a region level (9 studies; 0.70, 95% CI 0.53-0.83). Studies using T2WI as the only imaging modality had lower sensitivity (30 studies; 0.53, 95% CI 0.44-0.63) than studies using an additional functional technique (18 studies; 0.63, 95% CI 0.51-0.74). Furthermore, the 15 studies with a 3.0T device had higher sensitivity (0.61, 95% CI 0.48-0.72) but lower specificity (0.88, 95% CI 0.82-0.92) in comparison to studies with 1.0 or 1.5T (33 studies; sensitivity 0.55, 95% CI 0.45-0.65; specificity 0.92, 95% CI 0.88-0.94). Studies that used a 3.0T device and no endorectal coil (5 studies) showed the highest sensitivity (0.71, 95% CI 0.51-0.86), with a specificity of 0.90 (95% CI 0.72-0.97). We found comparable sensitivity and specificity between studies with and without an endorectal coil. In addition, studies with more or less than 50 participants, with an absolute ECE prevalence of more or less than 10%, and studies with or without high risk in the QUADAS-2 assessment showed comparable pooled sensitivity and specificity estimates.

Sensitivity analyses for SVI (stage T3b) detection

Figure 4b shows the results for SVI (T3b). Studies reporting on SVI that used an endorectal coil showed a sensitivity and specificity of 0.59 (95% CI 0.50-0.67) and 0.97 (95% CI 0.95-0.98), compared to 0.51 (95% CI 0.23-0.78) and 0.94 (95% CI 0.90-0.97) in studies without an endorectal coil, respectively. Studies that used 1 or 2 functional techniques in addition to T2WI, showed higher sensitivity values (15 studies; 0.64 (95% CI 0.48-0.76) than studies using T2WI only (22 studies; 0.53, 95% CI 0.39-0.67). Studies using a 1.0 or 1.5T device had a sensitivity and specificity (27 studies; 0.58, 95% CI 0.46-0.70; 0.97, 95% CI 0.95-0.98) that was comparable to studies with a 3.0T scanner (10 studies; 0.57, 95% CI 0.37-0.75; 0.95, 95% CI 0.91-0.97). When studies used a 3.0T scanner without endorectal coil (5 studies) the sensitivity was higher than with an endorectal coil; 0.65 (95% CI 0.30-0.89) and 0.45 (95% CI 0.30-0.60), respectively. Studies with a 1.5T scanner showed higher sensitivity with an endorectal coil (0.62, 95% CI 0.51-0.71) than studies without an endorectal coil (0.37, 95% CI 0.08-0.80). The combination of 3T and mp-MRI had the highest sensitivity of 0.73 (5 studies; 95% CI 0.45-0.90) and specificity of 0.95 (95% CI 0.89-0.98). Because of the limited number of studies per stratum, and the low number of true positive and false negative cases, several sensitivity analyses resulted in unstable pooled estimates. Therefore, we were unable to reliably compare studies with and without an absolute SVI prevalence of more than 10%, with and without 50 participants or more, patient versus region-level analysis, and with different risk populations.

ECE

	Sensitivity	Specificity
Overall	0.57 (0.49-0.64)	0.91 (0.88-0.93)
With erc	0.57 (0.48-0.65)	0.92 (0.88-0.94)
Without erc	0.58 (0.43-0.72)	0.88 (0.81-0.92)
1T and 1.5T	0.55 (0.45-0.65)	0.92 (0.88-0.94)
3T	0.61 (0.48-0.72)	0.88 (0.82-0.92)
T2WI only	0.53 (0.44-0.63)	0.91 (0.86-0.94)
T2WI + one functional technique	0.62 (0.48-0.74)	0.91 (0.86-0.95)
T2WI + two functional techniques	0.69 (0.39-0.88)	0.91 (0.75-0.97)
mpMRI	0.63 (0.51-0.74)	0.91 (0.86-0.94)
3T + erc	0.60 (0.40-0.78)	0.88 (0.82-0.92)
3T + no erc	0.61 (0.45-0.75)	0.87 (0.77-0.93)
1.5T + erc	0.55 (0.45-0.65)	0.93 (0.89-0.95)
1.5T + no erc	0.54 (0.28-0.78)	0.89 (0.78-0.95)
3T + mpMRI	0.68 (0.50-0.82)	0.86 (0.77-0.92)
3T + no mpMRI	0.53 (0.38-0.68)	0.89 (0.82-0.94)
1.5T + mpMRI	0.60 (0.43-0.75)	0.94 (0.89-0.97)
1.5T + no mpMRI	0.53 (0.40-0.65)	0.91 (0.85-0.94)
mpMRI + erc	0.59 (0.44-0.73)	0.91 (0.87-0.94)
mpMRI + no erc	0.71 (0.51-0.86)	0.90 (0.72-0.97)
no mpMRI + erc	0.55 (0.44-0.66)	0.92 (0.87-0.95)
no mpMRI + no erc	0.49 (0.32-0.66)	0.86 (0.80-0.90)
Patient only	0.55 (0.47-0.63)	0.91 (0.88-0.93)
Region only	0.70 (0.53-0.83)	0.89 (0.78-0.95)
Participants 50 or more	0.57 (0.48-0.66)	0.90 (0.86-0.93)
Participants less than 50	0.56 (0.42-0.70)	0.92 (0.88-0.95)
Prevalence 10 or more	0.58 (0.49-0.65)	0.90 (0.87-0.92)
Prevalence less than 10	0.49 (0.28-0.71)	0.97 (0.93-0.99)*
QUADAS high risk	0.56 (0.45-0.67)	0.92 (0.86-0.96)
QUADAS no high risk	0.57 (0.47-0.67)	0.90 (0.86-0.93)
Risk category high	0.63 (0.35-0.85)	0.91 (0.81-0.96)
Risk category mixed	0.57 (0.48-0.65)	0.88 (0.84-0.91)
Risk category low	0.33 (0.27-0.40)	0.92 (0.84-0.96)*
Risk category NR	0.67 (0.49-0.81)	0.96 (0.91-0.99)



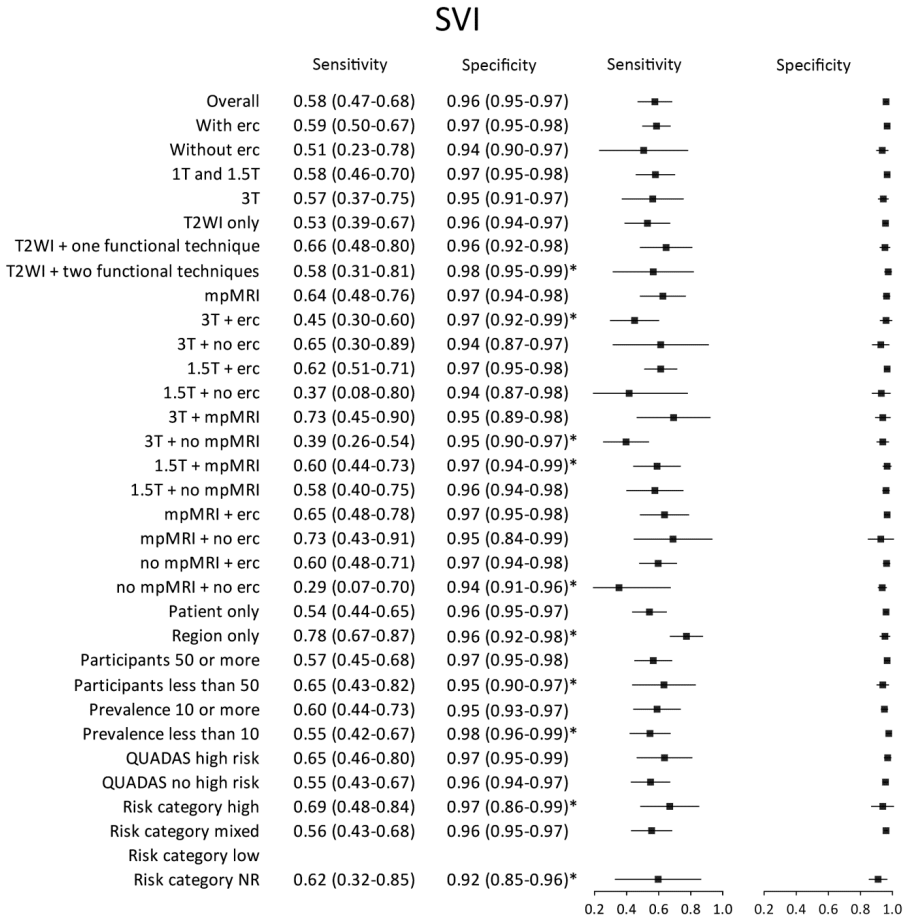


Figure 4a and b Forest plots of pooled sensitivity and specificity estimates, with corresponding 95% confidence intervals (CI) of all studies overall and for the different sensitivity analyses. Figure 4a provides the estimates for the assessment of extracapsular extension (ECE) and 4b for seminal vesicle invasion (SVI).
Abbreviations: erc = endorectal coil; mp-MRI = multiparametric MRI; NR = not reported T = Tesla; T2WI = T2 weighted imaging; QUADAS = quality assessment of diagnostic accuracy studies. * = unstable pooled estimates.

Sensitivity analyses for overall stage T3 detection

Supplementary figure 3 shows the results of studies that did not stratify between stage T3a and T3b, but reported on overall stage T3 only. Studies reporting on detection of overall stage T3 disease that used an endorectal coil had lower sensitivity (25 studies; 0.57, 95% CI 0.49-0.64) and higher specificity (0.90, 95% CI 0.86-0.93) than studies without an endorectal coil (13 studies; sensitivity 0.66, 95% CI 0.55-0.76; specificity 0.85, 95% CI 0.78-0.90). Use of a higher magnetic field strength of 3.0T yielded higher sensitivity (0.63, 95% CI 0.52-0.73) than 1.0T or 1.5T device (0.60, 95% CI 0.52-0.67).

Addition of functional imaging techniques to T2WI yielded higher sensitivity (0.62, 95% CI 0.52-0.71) and lower specificity (0.86, 95% CI 0.81-0.90) in comparison to studies that used T2WI alone (sensitivity 0.60, 95% CI 0.51-0.67, specificity 0.90, 95% CI 0.85-0.94). The absolute prevalence and the number of participants did not influence the sensitivity and specificity. We were unable to reliably compare studies with different risk populations, T3 prevalence of more or less than 10%, and patient versus region-level analysis.

Publication bias

The slope coefficients for Deeks funnel plots for detection of ECE, SVI, and overall stage T3 suggest symmetry in the data and a low likelihood of publication bias (ECE: p=0.88; SVI: p=0.88; T2 vs. T3: p=0.80).[86]

DISCUSSION

This meta-analysis of the diagnostic accuracy of MRI for local PCa staging revealed high specificity and poor and heterogeneous sensitivity overall. The pooled sensitivity and specificity for detection of ECE/T3a (45 studies, 5681 patients), SVI/T3b (34 studies, 5677 patients), and overall stage T3 disease (38 studies, 4001 patients) were 0.57 (95% CI 0.49-0.64) and 0.91 (95% CI 0.88-0.93), 0.58 (95% CI 0.47-0.68) and 0.96 (95% CI 0.95-0.97), and 0.61 (95% CI 0.54-0.67) and 0.88 (95% CI 0.85-0.91), respectively. Several of the patient, study, and imaging characteristics explored clearly influenced staging accuracy. This affected the sensitivity more, whereas specificity remained relatively stable.

In particular, use of functional techniques in addition to T2WI and of devices with a higher field strength appeared to have a large influence on sensitivity. Use of functional techniques is recommended and widely adopted for PCa detection, for example, the European Society of Urogenital Radiology consensus statement recommends the use of at least two functional techniques for detection of PCa.[87] For local staging, functional techniques with high resolution T2WI are also suggested to help radiologists to focus on lesions suspicious for local extension of the prostate.[88] This is in line with the current meta-analysis. When one additional functional technique (DWI, DCE-MRI, or MRSI) was used, sensitivity improved for ECE, SVI, and overall stage T3 compared to T2WI alone. When two or more functional techniques were used sensitivity further improved for ECE, but slightly decreased for overall T3 disease. For SVI detection, the pooled estimates for studies with two additional functional techniques were unstable.

Use of a higher field strength (3.0T instead of 1.0 or 1.5T) improved the detection sensitivity for ECE and overall stage T3. The use of an endorectal coil appeared useful for a field strength of 1.5T or in the absence of mp-MRI. However, when higher field strengths or additional functional techniques were used, studies that used an endorectal coil showed lower sensitivity than studies without an endorectal coil.

Region-based image interpretation showed higher sensitivity for local staging than patient-based analysis. This is possibly because this technique artificially increases the number of true positives, leading to inflated sensitivity. It is important to be aware of this effect when interpreting studies that report at a region instead of patient level.

The major strength of our meta-analysis is that it provides a complete and unique overview of the literature since the last extensive meta-analysis by Engelbrecht et al.,[89] and the query was not restricted to certain imaging parameters as in the meta-analysis by Silva et al.[7] Therefore, many studies could be included and sensitivity analyses of the most important patient, study, and imaging characteristics were possible. The outcomes of these sensitivity analyses could aid radiologists and urologists in deciding on which imaging protocols to use for local staging of PCa.

Some methodologic issues also need to be considered. First, we could not completely explain the heterogeneity, because many studies did not include sufficient information for all study characteristics. Information was often missing for blinding to clinical information while interpreting the index test, the risk profile of the study population, and image interpretation methods. In addition, some of the studied strata were too small to result in stable pooled estimates.

Second, because several studies included multiple analyses within the same patient group (eg, for different readers, functional techniques, or coils) we had to choose one set among the accuracy results presented. Several studies reported results for different experience levels.[19,31-33,38,47,64-66,69] We decided to include the most up--to-date technique, and preferred the most experienced radiologist. We believe this choice resembles clinical practice and demonstrates the potential benefit of the newest techniques. Even though this is the best approximation of the clinical situation, specificity results of the least experienced readers were comparable to the most experienced readers. However, the sensitivity values, with worse results for the less experienced readers, although comparable or even better results were also reported. Use of the results from the less experienced readers would have led to

only minor changes in the results, and would therefore not bias the overall conclusions of our meta-analysis.

Currently, MRI is the best imaging technique available for assessing extra-prostatic extension in clinical practice. PSA, DRE and transrectal ultrasound are not accurate enough for local staging (T stage), but other imaging techniques (e.g. PET-CT) can be of value for detection of lymph nodes or distant metastases. These techniques are not accurate enough to assess extra-prostatic extension.[2,90]

The studies included used different methods to standardise reporting; dichotomisation, Likert scales, or a standardized lexicon. Many studies did not sufficiently describe the reporting method, which precluded a sensitivity analysis regarding this characteristic. Wibmer et al. showed that a standardized reporting system using a 5-point Likert scale with a standardized lexicon, could improve staging accuracy over a nonstandardized approach.[91] In working towards a robust method, an international language should preferably be used, similar to the PI-RADS for PCa detection, to improve detection accuracy.[92]

MRI is limited for detection of focal (microscopic) ECE, a disease category with favourable prognosis compared to more extensive ECE.[93,94] Staging accuracy appears to decrease when cases of focal ECE are incorporated.[19,33] However, there is no internationally accepted definition of focal compared to established ECE.[95] Refinement is needed for both clinical/pathological and imaging criteria to further investigate the clinical value of MRI in detecting focal ECE.

The current meta-analysis shows that MRI has high specificity but low sensitivity. Traditionally, radiologists have focused on high-specificity reading in order to minimize unnecessary exclusion of men from curative treatment. This is probably why the meta-analysis revealed high specificity and low sensitivity for MRI. Nowadays, urologists become more interested in high-sensitivity reading, to reduce positive surgical margins and preserve neurovascular bundles. Our analyses show that a combination of high magnetic field strength (3.0T) and functional imaging techniques can slightly improve the sensitivity of MRI. However, on its own the technique is not good enough to accurately stage local PCa. Prediction of the correct T stage can improve when MRI findings are combined with clinical data such as D'Amico risk categories.[2] In the future, a risk-tailored approach might be more appropriate. Unfortunately, results for the sensitivity analyses for these risk groups could not be pooled due to unstable estimates, especially for the low-risk group. The trend

showed lower sensitivity values for ECE and SVI when the risk on extra-prostatic extension was low. A risk-tailored approach is therefore warranted whereby radiologists perform high specificity reading for high-risk patients to reduce the risk of positive surgical margins, and high-sensitivity reading for patients with low- to intermediate-risk to select candidates for curative treatment or active surveillance.[71]

CONCLUSIONS

MRI appears to have high specificity but poor and heterogeneous sensitivity for detection of ECE, SVI, and overall stage T3. Use of an endorectal coil yielded no additional benefit for ECE detection, but slightly improved the sensitivity of SVI detection. Higher field strengths and the use of additional functional imaging techniques seemed to improve accuracy of local staging.

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA: a cancer journal for clinicians* 2015.
2. Heidenreich A, Bastian PJ, Bellmunt J, et al. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *European urology* 2014; 65(1): 124-37.
3. Ward JF, Slezak JM, Blute ML, Bergstralh EJ, Zincke H. Radical prostatectomy for clinically advanced (cT3) prostate cancer since the advent of prostate-specific antigen testing: 15-year outcome. *BJU international* 2005; 95(6): 751-6.
4. Mullerad M, Hricak H, Kuroiwa K, et al. Comparison of endorectal magnetic resonance imaging, guided prostate biopsy and digital rectal examination in the preoperative anatomical localization of prostate cancer. *The Journal of urology* 2005; 174(6): 2158-63.
5. Moore CM, Robertson NL, Arsanious N, et al. Image-guided prostate biopsy using magnetic resonance imaging-derived targets: a systematic review. *European urology* 2013; 63(1): 125-40.
6. Schoots IG, Roobol MJ, Nieboer D, Bangma CH, Steyerberg EW, Hunink MG. Magnetic Resonance Imaging-targeted Biopsy May Enhance the Diagnostic Accuracy of Significant Prostate Cancer Detection Compared to Standard Transrectal Ultrasound-guided Biopsy: A Systematic Review and Meta-analysis. *European urology* 2014.
7. Silva RC, Sasse AD, Matheus WE, Ferreira U. Magnetic resonance image in the diagnosis and evaluation of extra-prostatic extension and involvement of seminal vesicles of prostate cancer: a systematic review of literature and meta-analysis. *International braz j urol : official journal of the Brazilian Society of Urology* 2013; 39(2): 155-66.
8. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine* 2009; 6(7): e1000097.
9. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine* 2011; 155(8): 529-36.
10. Takwoingi YD, J.J. MetaDAS: A SAS macro for meta-analysis of diagnostic accuracy studies. 2010.
11. Akin O, Agildere AM, Ersoy H, Ozen H, Tekin I, Ayhan A. Local staging of prostate cancer with endorectal surface coil MR imaging in a mid-field magnetic system. *Clinical imaging* 2003; 27(1): 47-51.
12. Akin O, Sala E, Moskowitz CS, et al. Transition zone prostate cancers: features, detection, localization, and staging at endorectal MR imaging. *Radiology* 2006; 239(3): 784-92.
13. Allen DJ, Hindley R, Clovis S, et al. Does body-coil magnetic-resonance imaging have a role in the preoperative staging of patients with clinically localized prostate cancer? *BJU international* 2004; 94(4): 534-8.
14. Armitage JN, Igali L, Ball RY, Mills RD, Kumar V. Comparison of body-array MRI and Partin tables for predicting extracapsular prostate cancer. *Journal of Clinical Urology* 2013; 6(1): 50-4.
15. Augustin H, Fritz GA, Ehammer T, Auprich M, Pummer K. Accuracy of 3-Tesla magnetic resonance imaging for the staging of prostate cancer in comparison to the Partin tables. *Acta radiologica (Stockholm, Sweden : 1987)* 2009; 50(5): 562-9.
16. Bernstein MR, Cangiano T, D'Amico A, et al. Endorectal coil magnetic resonance imaging and clinicopathologic findings in T1c adenocarcinoma of the prostate. *Urologic oncology* 2000; 5(3): 104-7.
17. Beyersdorff D, Taymoorian K, Knosel T, et al. MRI of prostate cancer at 1.5 and 3.0 T: comparison of image quality in tumor detection and staging. *AJR American journal of roentgenology* 2005; 185(5): 1214-20.
18. Bloch BN, Furman-Haran E, Helbich TH, et al. Prostate cancer: Accurate determination of extracapsular extension with high-spatial-resolution dynamic contrast-enhanced and T2-weighted MR imaging - Initial results. *Radiology* 2007; 245(1): 176-85.

19. Bloch BN, Genega EM, Costa DN, et al. Prediction of prostate cancer extracapsular extension with high spatial resolution dynamic contrast-enhanced 3-T MRI. *European radiology* 2012; 22(10): 2201-10.

20. Borre M, Lundorf E, Marcussen N, Langkilde NC, Wolf H. Phased array magnetic resonance imaging for staging clinically localized prostate cancer. *Acta Oncologica* 2005; 44(6): 589-92.

21. Brajtbord JS, Lavery HJ, Nabizada-Pace F, Senaratne P, Samadi DB. The ability of endorectal coil MRI to preoperatively predict seminal vesicle invasion. *International Journal of Medical Robotics and Computer Assisted Surgery* 2011; 7: 28-9.

22. Brassell SA, Krueger WR, Choi JH, Taylor JA, 3rd. Correlation of endorectal coil magnetic resonance imaging of the prostate with pathologic stage. *World J Urol* 2004; 22(4): 289-92.

23. Brown JA, Rodin DM, Harisinghani M, Dahl DM. Impact of preoperative endorectal MRI stage classification on neurovascular bundle sparing aggressiveness and the radical prostatectomy positive margin rate. *Urologic oncology* 2009; 27(2): 174-9.

24. Cerantola Y, Valerio M, Kawkabani Marchini A, Meuwly JY, Jichlinski P. Can 3T multiparametric magnetic resonance imaging accurately detect prostate cancer extracapsular extension? *Canadian Urological Association journal = Journal de l'Association des urologues du Canada* 2013; 7(11-12): E699-703.

25. Chandra RV, Heinze S, Dowling R, Shadbolt C, Costello A, Pedersen J. Endorectal magnetic resonance imaging staging of prostate cancer. *ANZ journal of surgery* 2007; 77(10): 860-5.

26. Chong Y, Kim CK, Park SY, Park BK, Kwon GY, Park JJ. Value of diffusion-weighted imaging at 3 T for prediction of extracapsular extension in patients with prostate cancer: a preliminary study. *AJR American journal of roentgenology* 2014; 202(4): 772-7.

27. Cornud F, Flam T, Chauveinc L, et al. Extraprostatic spread of clinically localized prostate cancer: factors predictive of pT3 tumor and of positive endorectal MR imaging examination results. *Radiology* 2002; 224(1): 203-10.

28. Cornud F, Hamida K, Flam T, et al. Endorectal color doppler sonography and endorectal MR imaging features of nonpalpable prostate cancer: correlation with radical prostatectomy findings. *AJR American journal of roentgenology* 2000; 175(4): 1161-8.

29. Cornud F, Rouanne M, Beuvon F, et al. Endorectal 3D T2-weighted 1mm-slice thickness MRI for prostate cancer staging at 1.5Tesla: should we reconsider the indirects signs of extracapsular extension according to the D'Amico tumor risk criteria? *European journal of radiology* 2012; 81(4): e591-7.

30. Counago F, Recio M, Del Cerro E, et al. Role of 3.0 T multiparametric MRI in local staging in prostate cancer and clinical implications for radiation oncology. *Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico* 2014.

31. Futterer JJ, Engelbrecht MR, Huisman HJ, et al. Staging prostate cancer with dynamic contrast-enhanced endorectal MR imaging prior to radical prostatectomy: experienced versus less experienced readers. *Radiology* 2005; 237(2): 541-9.

32. Futterer JJ, Engelbrecht MR, Jager GJ, et al. Prostate cancer: Comparison of local staging accuracy of pelvic phased-array coil alone versus integrated endorectal-pelvic phased-array coils. *Eur Radiol* 2007; 17(4): 1055-65.

33. Futterer JJ, Heijmink SW, Scheenen TW, et al. Prostate cancer: local staging at 3-T endorectal MR imaging--early experience. *Radiology* 2006; 238(1): 184-91.

34. Graser A, Heuck A, Sommer B, et al. Per-sextant localization and staging of prostate cancer: correlation of imaging findings with whole-mount step section histopathology. *AJR American journal of roentgenology* 2007; 188(1): 84-90.

35. Gupta RT, Faridi KF, Singh AA, et al. Comparing 3-T multiparametric MRI and the Partin tables to predict organ-confined prostate cancer after radical prostatectomy. *Urologic oncology* 2014.

36. Hara T, Nakanishi H, Nakagawa T, et al. Ability of preoperative 3.0-Tesla magnetic resonance imaging to predict the absence of side-specific extracapsular extension of prostate cancer. *International journal of urology : official journal of the Japanese Urological Association* 2013; 20(10): 993-9.

37. Hegde JV, Chen MH, Mulkern RV, Fennessy FM, D'Amico AV, Tempany CM. Preoperative 3-Tesla multiparametric endorectal magnetic resonance imaging findings and the odds of upgrading and upstaging at radical prostatectomy in men with clinically localized prostate cancer. *International journal of radiation oncology, biology, physics* 2013; 85(2): e101-7.

38. Heijmink SW, Futterer JJ, Hambrock T, et al. Prostate cancer: body-array versus endorectal coil MR imaging at 3 T--comparison of image quality, localization, and staging performance. *Radiology* 2007; 244(1): 184-95.

39. Hole KH, Axcrona K, Lie AK, et al. Routine pelvic MRI using phased-array coil for detection of extraprostatic tumour extension: accuracy and clinical significance. *Eur Radiol* 2013; 23(4): 1158-66.

40. Hwii Ko Y, Jae Sung D, Gu Kang S, et al. The predictability of T3 disease in staging MRI following prostate biopsy decreases in patients with high initial PSA and Gleason score. *Asian journal of andrology* 2011; 13(3): 487-93.

41. Ikonen S, Karkkainen P, Kivisaari L, et al. Endorectal magnetic resonance imaging of prostatic cancer: comparison between fat-suppressed T2-weighted fast spin echo and three-dimensional dual-echo, steady-state sequences. *Eur Radiol* 2001; 11(2): 236-41.

42. Jeong IG, Lim JH, You D, et al. Incremental value of magnetic resonance imaging for clinically high risk prostate cancer in 922 radical prostatectomies. *The Journal of urology* 2013; 190(6): 2054-60.

43. Johnston R, Wong LM, Warren A, Shah N, Neal D. The role of 1.5 Tesla magnetic resonance imaging in staging prostate cancer. *ANZ journal of surgery* 2013; 83(4): 234-8.

44. Jung DC, Lee HJ, Kim SH, Choe GY, Lee SE. Preoperative MR imaging in the evaluation of seminal vesicle invasion in prostate cancer: pattern analysis of seminal vesicle lesions. *Journal of magnetic resonance imaging : JMRI* 2008; 28(1): 144-50.

45. Kim B, Breau RH, Papadatos D, et al. Diagnostic accuracy of surface coil magnetic resonance imaging at 1.5 T for local staging of elevated risk prostate cancer. *Canadian Urological Association journal = Journal de l'Association des urologues du Canada* 2010; 4(4): 257-62.

46. Kim BS, Kim TH, Kwon TG, Yoo ES. Comparison of Pelvic Phased-Array versus Endorectal Coil Magnetic Resonance Imaging at 3 Tesla for Local Staging of Prostate Cancer. *Yonsei medical journal* 2012; 53(3): 550-6.

47. Kim CK, Choi D, Park BK, Kwon GY, Lim HK. Diffusion-weighted MR imaging for the evaluation of seminal vesicle invasion in prostate cancer: initial results. *Journal of magnetic resonance imaging : JMRI* 2008; 28(4): 963-9.

48. Kwek JW, Thng CH, Tan PH, et al. Phased-array magnetic resonance imaging of the prostate with correlation to radical prostatectomy specimens: local experience. *Asian journal of surgery / Asian Surgical Association* 2004; 27(3): 219-24; discussion 25-6.

49. Latchamsetty KC, Borden LS, Jr., Porter CR, et al. Experience improves staging accuracy of endorectal magnetic resonance imaging in prostate cancer: what is the learning curve? *The Canadian journal of urology* 2007; 14(1): 3429-34.

50. Lee SH, Park KK, Choi KH, et al. Is endorectal coil necessary for the staging of clinically localized prostate cancer? Comparison of non-endorectal versus endorectal MR imaging. *World J Urol* 2010; 28(6): 667-72.

51. May F, Treumann T, Dettmar P, Hartung R, Breul J. Limited value of endorectal magnetic resonance imaging and transrectal ultrasonography in the staging of clinically localized prostate cancer. *BJU international* 2001; 87(1): 66-9.

52. Min BD, Kim WT, Cho BS, et al. Usefulness of a combined approach of t1-weighted, t2-weighted, dynamic contrast-enhanced, and diffusion-weighted imaging in prostate cancer. *Korean journal of urology* 2012; 53(12): 830-5.

53. Nakashima J, Tanimoto A, Imai Y, et al. Endorectal MRI for prediction of tumor site, tumor size, and local extension of prostate cancer. *Urology* 2004; 64(1): 101-5.

54. Nepple KG, Rosevear HM, Stolpen AH, Brown JA, Williams RD. Concordance of preoperative prostate endorectal MRI with subsequent prostatectomy specimen in high-risk prostate cancer patients. *Urologic oncology* 2011.

55. Ogura K, Maekawa S, Okubo K, et al. Dynamic endorectal magnetic resonance imaging for local staging and detection of neurovascular bundle involvement of prostate cancer: correlation with histopathologic results. *Urology* 2001; 57(4): 721-6.

56. Oon SF, Power SP, Kelly JS, McDermott V, Ryan P, Ryan PC. The accuracy of magnetic resonance imaging in prostate cancer staging: a single-institution experience. *Irish journal of medical science* 2014.

57. Otto J, Thormer G, Seiwerts M, et al. Value of Endorectal Magnetic Resonance Imaging at 3T for the Local Staging of Prostate Cancer. *RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin* 2014; 186(8): 795-802.

58. Park BH, Choo SH, Jang HJ, et al. Accuracy of phased-array 3.0-T MR imaging in predicting extracapsular extension and influence on the decision to preserve neurovascular bundles at robotic assisted laparoscopic radical prostatectomy. *European Urology, Supplements* 2013; 12 (1): e297.

59. Park BK, Kim B, Kim CK, Lee HM, Kwon GY. Comparison of phased-array 3.0-T and endorectal 1.5-T magnetic resonance imaging in the evaluation of local staging accuracy for prostate cancer. *Journal of computer assisted tomography* 2007; 31(4): 534-8.

60. Park SY, Kim JJ, Kim TH, et al. The role of endorectal magnetic resonance imaging in predicting extraprostatic extension and seminal vesicle invasion in clinically localized prostate cancer. *Korean journal of urology* 2010; 51(5): 308-12.

61. Porcaro AB, Borsato A, Romano M, et al. Accuracy of preoperative endo-rectal coil magnetic resonance imaging in detecting clinical under-staging of localized prostate cancer. *World J Urol* 2013; 31(5): 1245-51.

62. Pugh TJ, Frank SJ, Achim M, et al. Endorectal magnetic resonance imaging for predicting pathologic T3 disease in Gleason score 7 prostate cancer: Implications for prostate brachytherapy. *Brachytherapy* 2013; 12(3): 204-9.

63. Ren J, Huan Y, Wang H, et al. Seminal vesicle invasion in prostate cancer: prediction with combined T2-weighted and diffusion-weighted MR imaging. *Eur Radiol* 2009; 19(10): 2481-6.

64. Renard-Penna R, Roupret M, Comperat E, et al. Accuracy of high resolution (1.5 tesla) pelvic phased array magnetic resonance imaging (MRI) in staging prostate cancer in candidates for radical prostatectomy: Results from a prospective study. *Urologic oncology* 2011.

65. Roethke M, Kaufmann S, Kniess M, et al. Seminal vesicle invasion: accuracy and analysis of infiltration patterns with high-spatial resolution T2-weighted sequences on endorectal magnetic resonance imaging. *Urologia internationalis* 2014; 92(3): 294-9.

66. Roethke MC, Lichy MP, Kniess M, et al. Accuracy of preoperative endorectal MRI in predicting extracapsular extension and influence on neurovascular bundle sparing in radical prostatectomy. *World J Urol* 2013; 31(5): 1111-6.

67. Rosenkrantz AB, Chandarana H, Gilet A, et al. Prostate cancer: utility of diffusion-weighted imaging as a marker of side-specific risk of extracapsular extension. *Journal of magnetic resonance imaging : JMRI* 2013; 38(2): 312-9.

68. Rud E, Klotz D, Rennesund K, et al. Preoperative magnetic resonance imaging for detecting uni- and bilateral extraprostatic disease in patients with prostate cancer. *World J Urol* 2014.

69. Ruprecht O, Weisser P, Bodelle B, Ackermann H, Vogl TJ. MRI of the prostate: interobserver agreement compared with histopathologic outcome after radical prostatectomy. *European journal of radiology* 2012; 81(3): 456-60.

70. Sala E, Akin O, Moskowitz CS, et al. Endorectal MR imaging in the evaluation of seminal vesicle invasion: diagnostic accuracy and multivariate feature analysis. *Radiology* 2006; 238(3): 929-37.

71. Somford DM, Hamoen EH, Futterer JJ, et al. The predictive value of endorectal 3 Tesla multiparametric magnetic resonance imaging for extraprostatic extension in patients with low, intermediate and high risk prostate cancer. *The Journal of urology* 2013; 190(5): 1728-34.

72. Soulie M, Aziza R, Escourrou G, et al. Assessment of the risk of positive surgical margins with pelvic phased-array magnetic resonance imaging in patients with clinically localized prostate cancer: a prospective study. *Urology* 2001; 58(2): 228-32.

73. Soylu FN, Peng Y, Jiang Y, et al. Seminal vesicle invasion in prostate cancer: evaluation by using multiparametric endorectal MR imaging. *Radiology* 2013; 267(3): 797-806.

74. Tan JS, Thng CH, Tan PH, et al. Local experience of endorectal magnetic resonance imaging of prostate with correlation to radical prostatectomy specimens. *Annals of the Academy of Medicine, Singapore* 2008; 37(1): 40-3.

75. Tanaka K, Shigemura K, Muramaki M, Takahashi S, Miyake H, Fujisawa M. Efficacy of using three-tesla magnetic resonance imaging diagnosis of capsule invasion for decision-making about neurovascular bundle preservation in robotic-assisted radical prostatectomy. *Korean journal of urology* 2013; 54(7): 437-41.

76. Torricelli P, Barberini A, Cinquantini F, Sighinolfi M, Cesinaro AM. 3-T MRI with phased-array coil in local staging of prostatic cancer. *Academic radiology* 2008; 15(9): 1118-25.

77. Torricelli P, Cinquantini F, Ligabue G, Bianchi G, Sighinolfi P, Romagnoli R. Comparative evaluation between external phased array coil at 3 T and endorectal coil at 1.5 T: preliminary results. *Journal of computer assisted tomography* 2006; 30(3): 355-61.

78. Tsao CW, Lin MH, Wu ST, et al. Combining prostrate-specific antigen and Gleason score increases the diagnostic power of endorectal coil magnetic resonance imaging in prostate cancer pathological stage. *Journal of the Chinese Medical Association : JCMA* 2013; 76(1): 20-4.

79. Wang JG, Huang J, Chin AI. RARP in high-risk prostate cancer: use of multi-parametric MRI and nerve sparing techniques. *Asian journal of andrology* 2014.

80. Wang L, Mullerad M, Chen HN, et al. Prostate cancer: incremental value of endorectal MR imaging findings for prediction of extracapsular extension. *Radiology* 2004; 232(1): 133-9.

81. Wang L, Zhang J, Schwartz LH, et al. Incremental value of multiplanar cross-referencing for prostate cancer staging with endorectal MRI. *AJR American journal of roentgenology* 2007; 188(1): 99-104.

82. Wetter A, Engl TA, Nadjmabadi D, et al. Combined MRI and MR spectroscopy of the prostate before radical prostatectomy. *AJR American journal of roentgenology* 2006; 187(3): 724-30.

83. Xylinas E, Yates DR, Renard-Penna R, et al. Role of pelvic phased array magnetic resonance imaging in staging of prostate cancer specifically in patients diagnosed with clinically locally advanced tumours by digital rectal examination. *World J Urol* 2011.

84. Zhang J, Hricak H, Shukla-Dave A, et al. Clinical stage T1c prostate cancer: evaluation with endorectal MR imaging and MR spectroscopic imaging. *Radiology* 2009; 253(2): 425-34.

85. Zhang JQ, Loughlin KR, Zou KH, Haker S, Tempany CM. Role of endorectal coil magnetic resonance imaging in treatment of patients with prostate cancer and in determining radical prostatectomy surgical margin status: report of a single surgeon's practice. *Urology* 2007; 69(6): 1134-7.

86. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of clinical epidemiology* 2005; 58(9): 882-93.

87. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol* 2012; 22(4): 746-57.

88. Futterer JJ, Barentsz JO, Heijmink SW. Value of 3-T magnetic resonance imaging in local staging of prostate cancer. *Topics in magnetic resonance imaging : TMRI* 2008; 19(6): 285-9.

89. Engelbrecht MR, Jager GJ, Laheij RJ, Verbeek AL, van Lier HJ, Barentsz JO. Local staging of prostate cancer using magnetic resonance imaging: a meta-analysis. *European radiology* 2002; 12(9): 2294-302.

90. Umbehr MH, Muntener M, Hany T, Sulser T, Bachmann LM. The role of 11C-choline and 18F-fluorocholine positron emission tomography (PET) and PET/CT in prostate cancer: a systematic review and meta-analysis. *European urology* 2013; 64(1): 106-17.

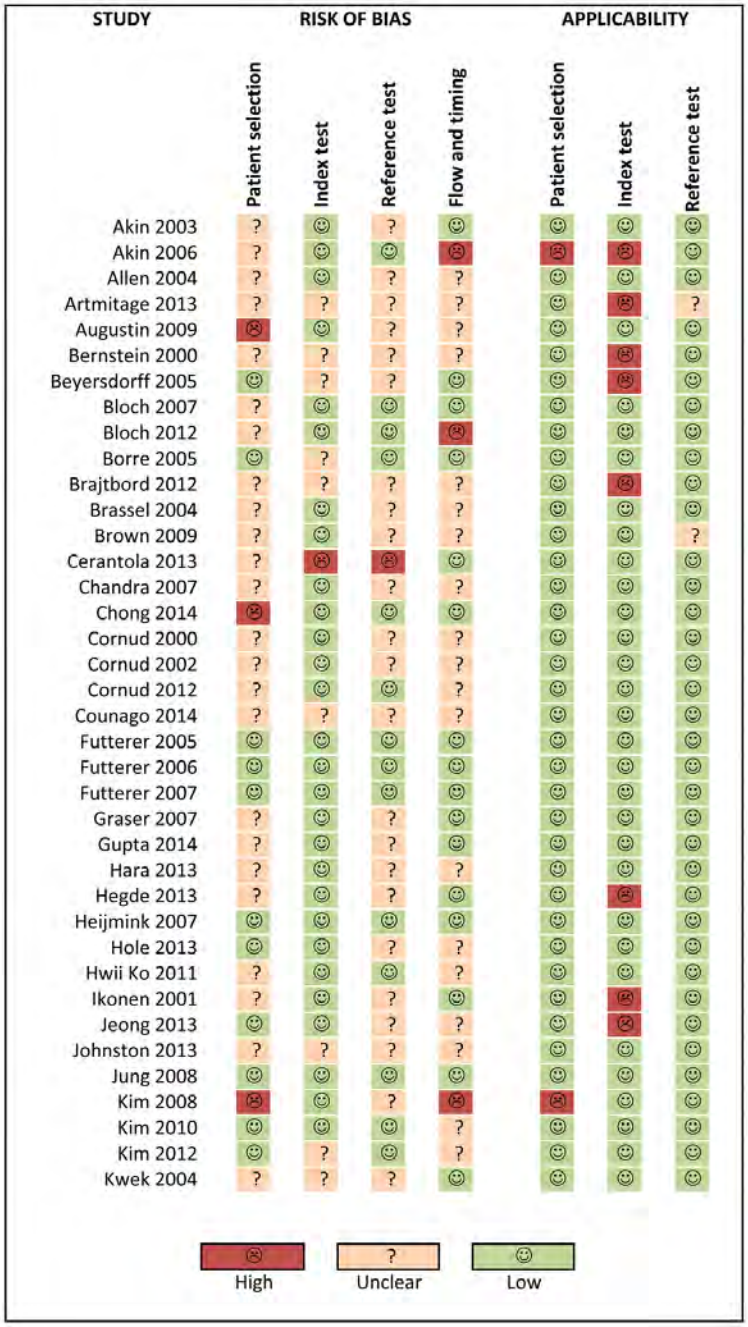
91. Wibmer A, Vargas HA, Sosa R, Zheng J, Moskowitz C, Hricak H. Value of a standardized lexicon for reporting levels of diagnostic certainty in prostate MRI. *AJR American journal of roentgenology* 2014; 203(6): W651-7.

92. Hamoen EH, de Rooij M, Witjes JA, Barentsz JO, Rovers MM. Use of the Prostate Imaging Reporting and Data System (PI-RADS) for Prostate Cancer Detection with Multiparametric Magnetic Resonance Imaging: A Diagnostic Meta-analysis. *European urology* 2015; 67(6): 1112-21.

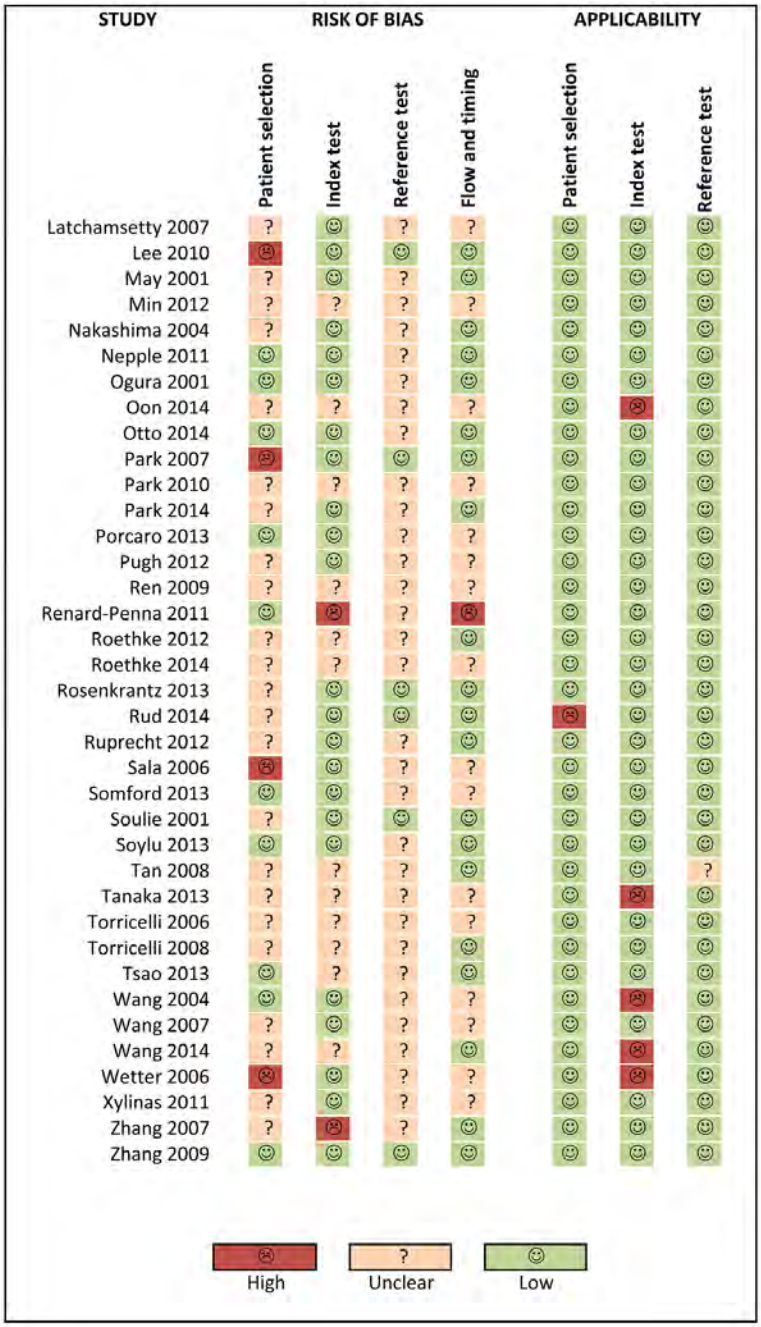
93. Chuang AY, Nielsen ME, Hernandez DJ, Walsh PC, Epstein JI. The significance of positive surgical margin in areas of capsular incision in otherwise organ confined disease at radical prostatectomy. *The Journal of urology* 2007; 178(4 Pt 1): 1306-10.

94. Wheeler TM, Dillioglugil O, Kattan MW, et al. Clinical and pathological significance of the level and extent of capsular invasion in clinical stage T1-2 prostate cancer. *Human pathology* 1998; 29(8): 856-62..

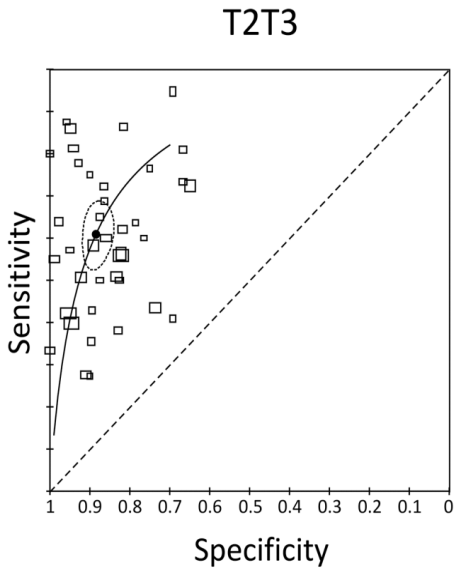
95. Fine SW, Amin MB, Berney DM, et al. A contemporary update on pathology reporting for prostate cancer: biopsy and radical prostatectomy specimens. *European urology* 2012; 62(1): 20-39.



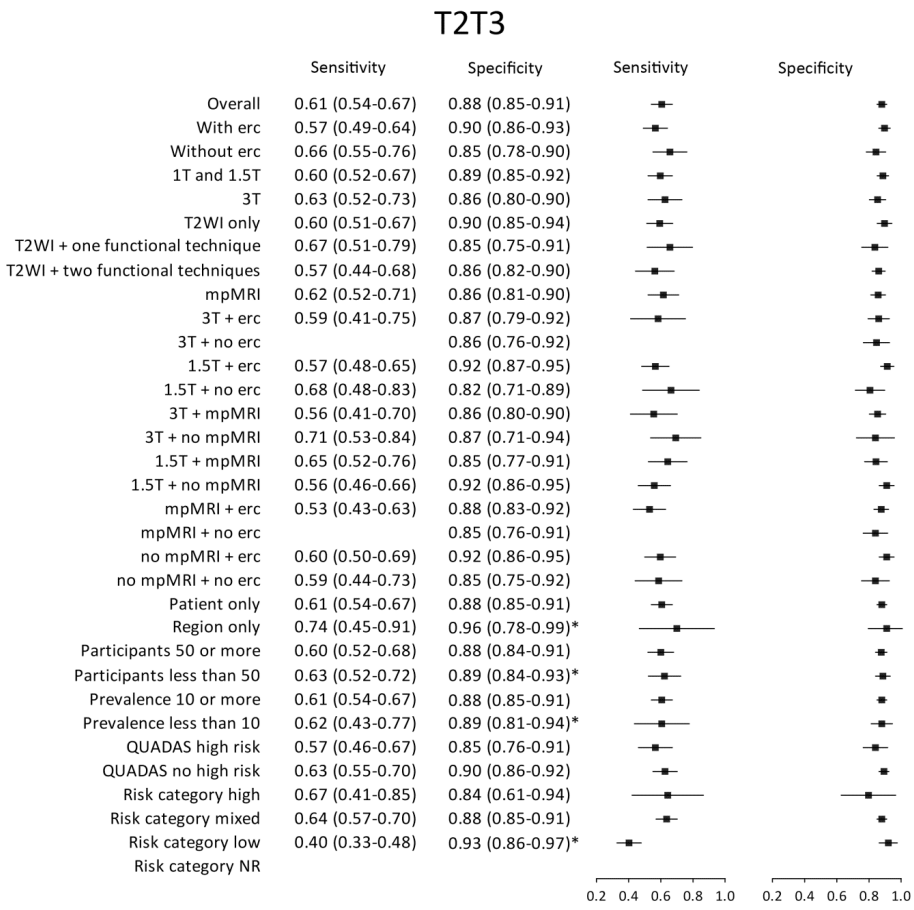
Supplemental figure 1 Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.



Supplemental figure 1 (continued) Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.



Supplemental figure 2 Hierarchic summary receiver operating characteristic (HSROC) plot (solid line) with summary point and 95% confidence interval (CI) area (circled area) for overall stage T3. Dashed line = no discrimination line (AUC of 0.5, meaning a worthless test), □ =data from individual studies included in meta-analysis for overall stage T3 disease (T2T3). The size of the squares corresponds with the relative size of the study population.



Supplemental figure 3 Forest plots of pooled sensitivity and specificity estimates for overall stage T3, with corresponding 95% confidence intervals (CIs) of all studies overall and for the different sensitivity analyses.

Abbreviations: erc = endorectal coil; mp-MRI = multiparametric MRI; T = Tesla; T2WI = T2 weighted imaging; QUADAS = quality assessment of diagnostic accuracy studies. * = unstable pooled estimates.


PART TWO
Chapters 5, 6
Costs and cost-effectiveness

5

Cost-effectiveness of MRI and MR-guided targeted biopsy versus systematic TRUS-GB in diagnosing prostate cancer: a modelling study from a health care perspective

Published as

de Rooij M, Crienen S, Witjes JA, Barentsz JO, Rovers MM, Grutters JPC

Cost-effectiveness of MRI and MR-guided targeted biopsy versus systematic TRUS-guided biopsy in diagnosing prostate cancer: a modelling study from a health care perspective

European Urology 2014 Sep;66(3):430-6

ABSTRACT

Background and objective

The current diagnostic strategy using transrectal ultrasound guided biopsy (TRUS-GB) raises concerns regarding over-diagnosis and over-treatment of prostate cancer (PCa). Interest in integrating multiparametric MRI (mp-MRI) and MR-guided biopsy (MRGB) into the diagnostic pathway to reduce over-diagnosis and improve grading is gaining ground, but it remains uncertain whether this image based strategy is cost-effective. The objective was to determine the cost-effectiveness of mp-MRI and MRGB compared with TRUS-GB.

Design, setting, and participants

A combined decision tree and Markov model for men with elevated PSA (> 4 ng/mL) was developed. Input data were derived from systematic literature searches, meta-analyses, and expert opinion.

Outcome measurements and statistical analysis

Quality-adjusted life years (QALYs) and health care costs of both strategies were modelled over 10 years after initial suspicion of PCa. Probabilistic and threshold analyses were performed to assess uncertainty.

Results and limitations

Despite uncertainty around the presented cost-effectiveness estimates, our results suggest that the MRI strategy is cost-effective compared with the standard of care. Expected costs per patient were €2423 for the MRI strategy and €2392 for the TRUS-GB strategy. Corresponding QALYs were higher for the MRI strategy (7.00 versus 6.90), resulting in an incremental cost-effectiveness ratio of €323/QALY. Threshold analysis revealed that the MRI strategy is cost-effective when sensitivity of MRGB is 20% or higher. The probability that the MRI strategy is cost-effective is around 80% at willingness to pay thresholds over €2000/QALY.

Conclusions

Total costs of the MRI strategy are almost equal with the standard of care, while reduction of over-diagnosis and over-treatment with the MRI strategy leads to an improvement in quality of life.

Patient summary

We compared costs and quality of life of the standard 'blind' diagnostic technique with an image-based technique for men with suspicion of prostate cancer. Our results suggest that costs were comparable, with higher quality of life for the image based technique.

INTRODUCTION

Systematic 10- to 12-core transrectal ultrasound guided biopsy (TRUS-GB) is the most accepted method for making a definite diagnosis of prostate cancer (PCa) in men with an increased serum prostate specific antigen (PSA) or abnormal digital rectal examination (DRE). Although advances have been made since the average number of TRUS-GB cores has increased from 6 to 12,[1] the probability of detecting PCa is still subject to random error because the operator cannot reliably visualize tumour.

The current TRUS-GB-diagnostic pathway is limited because of over-diagnosis and subsequent over-treatment of PCa[2-4] and is accompanied by a risk of post-biopsy infection.[5,6] This limitation may in turn lead to elevated psychological, clinical, and economic impacts[7,8] – the main reason the U.S. Preventive Services Task Force alerted the medical community to the dangers of PSA testing.[9] Therefore, the potential benefits of diagnosing PCa must be weighed against the risks, inforcing the need for better pretreatment characterization of PCa.

Multiparametric magnetic resonance imaging (mp-MRI) has emerged as an imaging technique that has the ability to accurately characterize PCa. This technology has led to opportunities to improve the diagnostic pathway.[10,11] With mp-MRI and subsequently MR-guided biopsy (MRGB), cancer-suspicious areas can be targeted.[12] As with mp-MRI, predominantly significant PCa is seen and insignificant cancer is not diagnosed, so this technique has the potential to solve the problem of over-diagnosis and over-treatment of the current TRUS-GB-pathway.[13,14] In addition, MRGB confers the ability to reduce unnecessary prostate biopsies by approximately 30-60%[15,16] using fewer biopsy cores (2-4 vs. 10-12).[16]

Prospective trials are currently performed to determine the definite diagnostic role of mp-MRI and MRGB compared with the current standard of TRUS-GB, but the decision regarding which diagnostic strategy to use should not be based on diagnostic accuracy alone. Costs related to performance and the therapeutic consequences of the test should also be taken into consideration.[17] In addition, it is important to look at other (in)direct consequences, such as quality of life (QoL) and survival. We therefore developed a decision-analytic model to assess from a healthcare perspective the cost-effectiveness of the MRI strategy (mp-MRI followed by MRGB) versus the standard TRUS-GB strategy in diagnosing PCa.

MATERIALS AND METHODS

Model development

We developed a decision analytic model to evaluate diagnostic accuracy, QoL, survival, and costs associated with two strategies for diagnosing PCa in patients with an elevated PSA level (>4 ng/mL). The model consisted of a decision tree combined with a Markov model (see Supplemental Figure 1, available online). Based on published clinical guidelines and expert opinion, a typical clinical setting was created. The first strategy is the current standard of care, where an elevated serum PSA is followed by systematic TRUS-GB. In this strategy, in case of a negative biopsy, the patient is followed-up by his urologist, with annual PSA testing, rectal examination and, when clinical suspicion remains, repeated systematic TRUS-GB. When a significant tumour is detected, the patient undergoes radical prostatectomy, radiotherapy, or watchful waiting. In the case of an insignificant tumour, the patient will undergo conservative treatment (active surveillance or watchful waiting) or a more invasive therapy, such as radiotherapy or radical prostatectomy. Clinical significance is based on clinical parameters (elevated serum PSA, DRE) and biopsy results (tumour aggressiveness in Gleason score and cancer core length) according to commonly used risk classification systems in PCa guidelines.[18,19] In this study, a Gleason score of 3+3 or a small-size 3+4 tumour is classified as insignificant, while large tumours with a Gleason score of 3+3 or tumours with ≥3+4 are classified as significant.

The second strategy is the experimental strategy, in which an elevated serum PSA is followed by mp-MRI. When a tumour-suspicious area is identified on mp-MRI, the patient will be scheduled for an MRGB. During the MRGB, the mp-MRI protocol is partially repeated to identify and target the previously reported suspicious area. In the case of a negative mp-MRI, the patient's urologists periodically follows up with him. When MRGB determines an insignificant or significant tumour, the patient undergoes the same treatment options as in the TRUS-GB strategy. We assumed that in patients who have a false-negative test result, the tumour would eventually be detected and treated.

To extrapolate the results, the decision tree was followed by a Markov model. Based on typical annual follow-up, a cycle time of 1 year was chosen with a 10-year time horizon, because after this period, no differences were expected between the strategies. To reflect the present value of the stream of costs and effects accruing over the time horizon of the analysis, quality adjusted life years (QALYs) were discounted by 1.5%, and costs were discounted by 4% according to Dutch guidelines.[20] In the Markov model, two main health states were defined based on whether patients were alive or dead.

Transition probabilities

The independent transition probabilities[12,13,21,22] and diagnostic accuracy data used for the TRUS-GB strategy were derived from systematic review of the literature[12,23-26] (Table 1). The accuracy data from the mp-MRI are based on the values estimated through a meta-analysis.[27] In the base case analysis, we assumed MRGB was 100% specific, with a sensitivity of 90% for targeting the tumour-suspicious regions. Patients were subdivided by their initial treatment: radical prostatectomy, radiotherapy (including brachytherapy), active surveillance, or watchful waiting using data from literature[28,29] and expert opinion.

Cost information

We used a health care perspective including only health care costs (Table 2). Unit costs were based on Dutch guidelines.[30] Heterogeneity between the different treatment options was taken into account by calculating the weighted mean of treatment alternatives (eg, for prostatectomy: 50% open and 50% robot-assisted laparoscopic approach). Costs of biopsy needles and complications were not taken into account. No difference was expected between the strategies in later costs, so no costs were assigned to the health states except for follow-up costs.

Outcome measures

Effectiveness was measured in terms of QALYs, which are the product of survival rates and the QoL (utility) associated with a health state. Survival[31-33] and utilities[34] were obtained from literature (Table 3). We assumed different survival for significant and insignificant tumours.[31] For the false-negatives, we assigned a utility based on the average of the treatment-dependent utilities because we assumed that the tumour would eventually be found and treated. Undetected significant tumours were assigned lower survival, because they receive treatment in a later, more advanced disease stage.[34]

Table 1 Probabilities used in decision analytic model.

Transition probabilities	P*	Source
Independent probabilities		
Tumour present when PSA elevated (4-6 ng/ml)	0.25 [SD=0.02]	Kranse et al. (2008)[22]
Tumour significant when present	0.50	Base assumption
Prostatectomy with significant tumour	0.40	Cooperberg et al. (2010)[29], Tewari et al. (2004)[28]
Radiotherapy with significant tumour	0.25	Cooperberg et al. (2010)[29], Tewari et al. (2004)[28]
Prostatectomy with insignificant tumour	0.10	Expert opinion
Brachytherapy with insignificant tumour	0.10	Expert opinion
Probabilities TRUS-GB strategy		
Sensitivity TRUS-GB	0.456 [α=221 β=266]	Hoeks et al. (2012)[12], Roethke et al. (2012)[25], France et al. (2011)[24], Hambrock et al. (2010)[23]
Specificity TRUS-GB	0.88 [α=56 β=8]	Taira et al. (2013)[26]
Correct estimation of tumour aggressiveness with TRUS-GB	0.53 [α=591 β=525]	Kvale et al (2009)[21]
Probabilities MRI strategy		
Sensitivity mp-MRI	0.74 [SD=0.06]	de Rooij et al. (2012)[27]
Specificity mp-MRI	0.88 [SD=0.05]	de Rooij et al. (2012)[27]
Sensitivity MRGB	0.90 [SD=0.05]	Base assumption
Specificity MRGB	1	Base assumption
Correct estimation of tumour aggressiveness with MRGB	0.88 [α=30 β=4]	Hambrock et al. (2012)[13]

Abbreviations: mp-MRI = multiparametric magnetic resonance imaging; MRGB = targeted magnetic resonance guided biopsies; PSA = prostate specific antigen; TRUS-GB = transrectal ultrasound guided biopsies; SD = standard deviation; β = beta value in the beta distribution; α = alpha value in the beta distribution. *Beta distributions were assigned to some of the parameters for use in the probabilistic sensitivity analysis. The characteristics of the beta distribution are presented between brackets, either as a standard deviation or as an alpha and beta value (where alpha represents the number of events in a sample, and beta the number of non-events).

Table 2 Cost data used in decision analytic model.

Description	Unit Costs (€)
Diagnostic procedure	
TRUS-GB	300
mp-MRI	345
MRGB	800
Histopathology analysis	231
Treatment	
Radical prostatectomy	12800
Radiation therapy	Insignificant tumour = 2401 Significant tumour = 4035
Watchful waiting/ active surveillance	100 (per year)

Abbreviations: mp-MRI = multiparametric magnetic resonance imaging; MRGB = targeted magnetic resonance guided biopsies; TRUS-GB = transrectal ultrasound guided biopsies.

Table 3 Utilities for health states used in decision analytic model.

Health State	Utility	SD*
Radical prostatectomy	0.67	0.29
Radiation therapy	0.73	0.30
Watchful waiting/ active surveillance	0.84	0.19

*All distributions are β distributions based on the mean value and standard deviation (SD).

Analysis

Baseline values were incorporated in the decision analytic model by using software (DATA, version Pro 2012; Tree Age Software, Williamstown, Massachusetts, USA). Incremental cost-effectiveness ratios (ICERs) were calculated by dividing the estimated difference in costs by the difference in QALYs, ie, costs per QALY gained. Whether the MRI strategy was deemed cost-effective depends on how much society is willing to pay for a QALY. If the ICER is lower than this willingness to pay (WTP), the MRI strategy is deemed cost-effective. In addition, diagnostic and treatment costs were calculated separately.

Threshold analysis was performed on the sensitivity of MRGB (base case 90%), the percentage of tumours (base case 25%), and the percentage of significant tumours (50%). Incremental net monetary benefits (iNMB) were calculated to present the results of the threshold analyses. To investigate sampling uncertainty concerning parameters in the model, probabilistic sensitivity analysis with 10,000 simulations was performed.[35]

Distributions were estimated for all uncertain parameters in the model except for the percentage of significant tumours and costs (Table 1). Parameters were assumed to be unrelated to each other. Results of the simulations are presented in cost-effectiveness planes and acceptability curves (CEACs).[36]

RESULTS

Cost-effectiveness of MRI and MRGB compared with TRUS-GB

The results show that the expected costs of the MRI strategy (€2423; 95% confidence interval (CI) €2219-€2637) were €31 higher than those for the TRUS-GB strategy (€2392; 95% CI €2227-€2563) (Table 4). The corresponding QALYs were 0.10 higher for the MRI strategy (7.00; 95% CI 3.72-8.32) compared with the TRUS-GB strategy (6.90; 95% CI 3.84-8.22) but with considerable uncertainty in these findings, as the CIs reflect. This resulted in an ICER of €323 per QALY gained.

Sensitivity analysis

The probability that the MRI strategy is more effective than the TRUS-GB strategy is 80%. The probability that the MRI strategy is both more effective and less costly is 25% (Figure 1). At WTP values of €1000/QALY and above, MRI becomes the strategy most likely to be cost-effective (Figure 2). At WTP values of €10000 and above, the MRI strategy is around 80% likely to be cost-effective.

The costs in the diagnostic and treatment pathway separately show that differences in costs are mainly generated in the treatment pathway (Figure 3a and 3b).

Threshold analysis shows that variation in sensitivity of MRGB affects the cost-effectiveness. The MRI strategy is cost-effective from a sensitivity threshold for MRGB of 20% (Figure 4). Varying the percentages of tumours and significant tumours was also found to change costs and effects (Figure 4), but for all values of these parameters, the MRI strategy was cost-effective compared with the TRUS-GB strategy.

Table 4 Baseline results from the analysis cost per QALY.

Strategy	Mean cost per strategy in € (95% CI)	Incremental costs in € (95% CI)	Effectiveness in QALY (95% CI)	Incremental QALYs (95% CI)	ICER (cost/ QALY)
TRUS-GB	2392 (22272563)		6.90 (3.84-8.22)		
MRI	2423 (22192637)	31 (-95162)	7.00 (3.72-8.32)	0.10 (-0.180.34)	323

Abbreviations: CI = confidence interval; ICER = incremental cost-effectiveness ratio; MRI = strategy with multiparametric magnetic resonance imaging and targeted magnetic resonance guided biopsies; QALY = quality-adjusted life year; TRUS-GB = strategy with transrectal ultrasound guided biopsy.

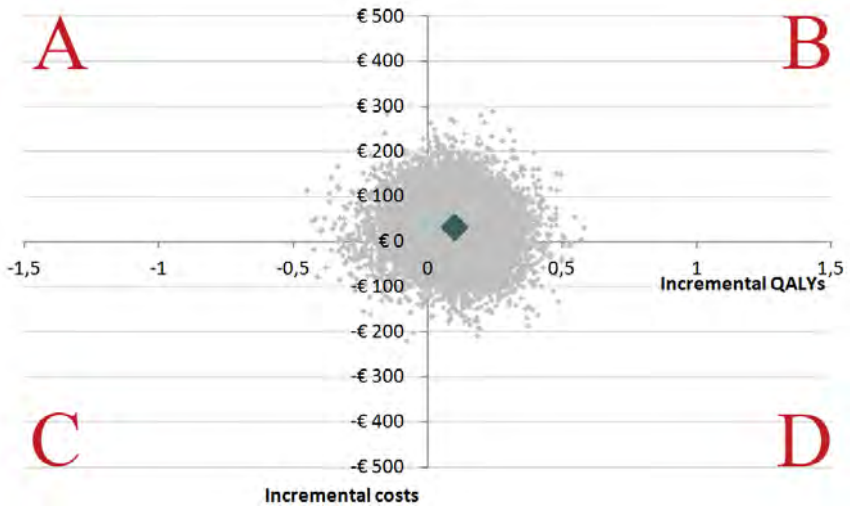


Figure 1 Scatter plot of probabilistic sensitivity analysis for MRI versus TRUS-GB strategies, including the total costs. A = MRI strategy less effective and more expensive; B = MRI strategy more effective and more expensive; C = MRI strategy less effective and less expensive; D = MRI strategy more effective and less expensive. The summary point (diamond) shows that the expected total costs of the MRI strategy are €31 higher than those for the TRUS-GB strategy, while the corresponding Quality Adjusted Life Years (QALYs) are 0.10 higher for the MRI strategy. The probability that the MRI strategy is more effective than the TRUS-GB strategy is 80% (B+D). The probability the MRI strategy is both more effective and less costly is 25% (D). Abbreviations: MRI = strategy with multiparametric magnetic resonance imaging and targeted magnetic resonance guided biopsies; QALY = quality-adjusted life year; TRUS-GB = strategy with transrectal ultrasound guided biopsy.

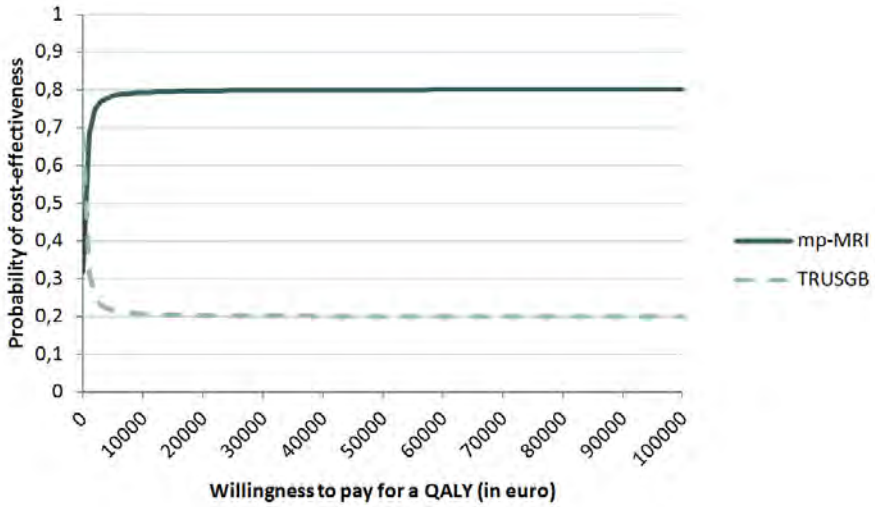


Figure 2 Cost-effectiveness acceptability curves, showing the probability that each strategy is the most cost-effective for a range of values of willingness to pay (WTP) per Quality Adjusted Life Year (QALY). For a WTP for the gain of a QALY of zero, the probability of MRI as the optimal strategy is 32. At WTP values of €1,000/QALY and above, MRI becomes the strategy that is most likely to be cost-effective. At WTP values of €10,000 and above, the MRI strategy is around 80% likely to be cost-effective. Abbreviations: MRI = strategy with multiparametric magnetic resonance imaging and targeted magnetic resonance guided biopsies; TRUS-GB = strategy with transrectal ultrasound guided biopsy.

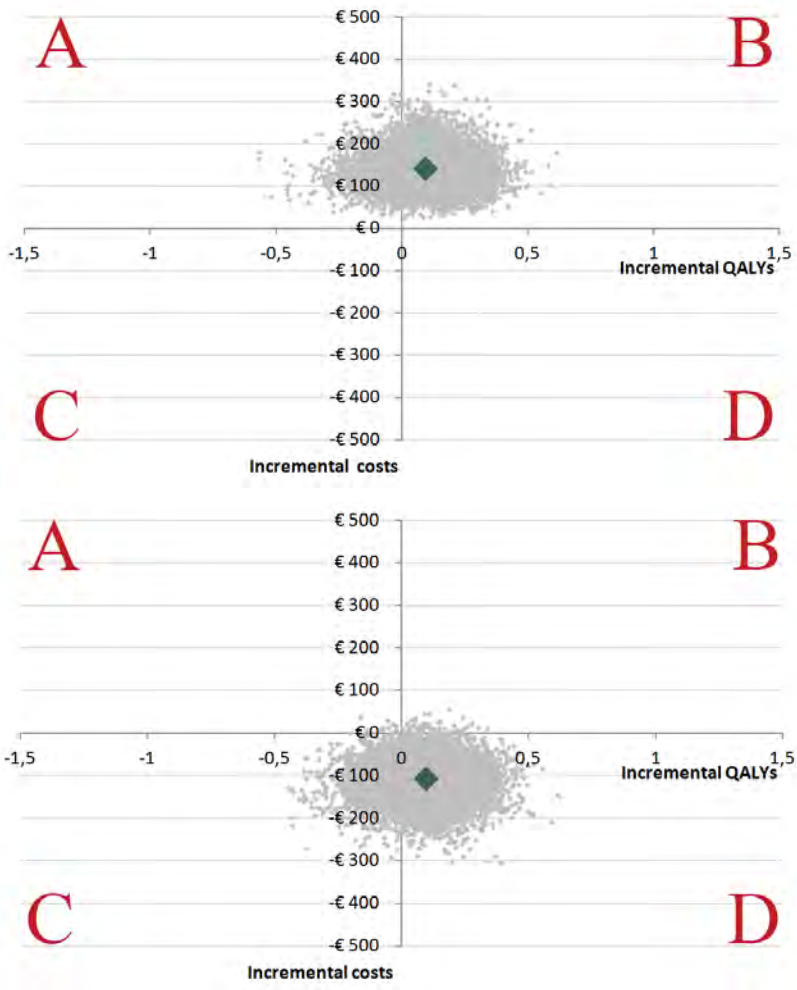


Figure 3a/b Scatter plot of probabilistic sensitivity analysis for MRI versus TRUS-GB strategies, analysing diagnostic costs (upper) and treatment cost (lower) only. A = MRI strategy less effective and more expensive; B = MRI strategy more effective and more expensive; C = MRI strategy less effective and less expensive; D = MRI strategy more effective and less expensive. Upper: The summary point (diamond) shows that the expected treatment costs of the MRI strategy are €140 higher than those for the TRUS-GB strategy, while the corresponding Quality Adjusted Life Years (QALYs) are 0.10 higher for the MRI strategy.
Lower: The summary point (diamond) shows that the expected diagnostic costs of the MRI strategy are €109 lower than those for the TRUS-GB strategy, while the corresponding Quality Adjusted Life Years (QALYs) are 0.10 higher for the MRI strategy.
Abbreviations: MRI = strategy with multiparametric magnetic resonance imaging and targeted magnetic resonance guided biopsies; QALY = quality-adjusted life year; TRUS-GB = strategy with random transrectal ultrasound guided biopsy.

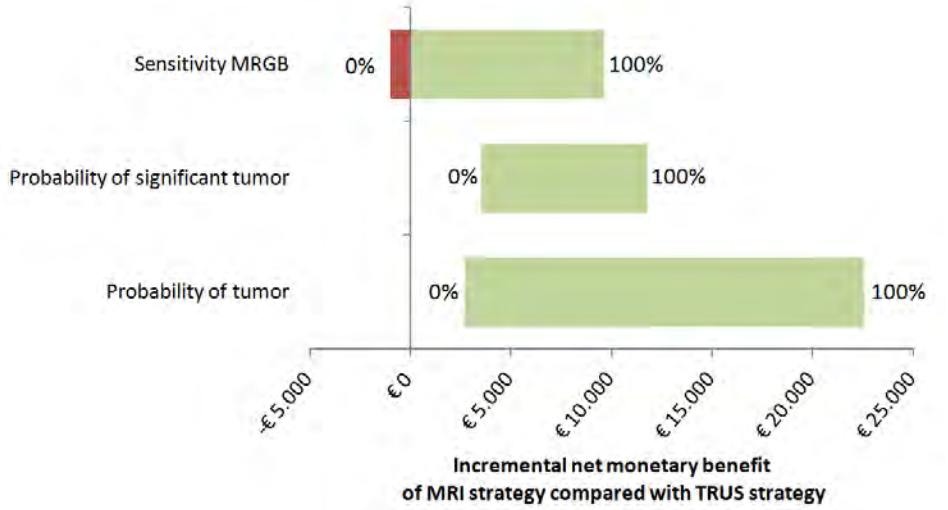


Figure 4 Incremental net monetary benefit (iNMB) for different values for sensitivity of MR guided biopsy (MRGB), probability of significant tumour, and probability of tumour in men with an elevated PSA. For all three parameters, values range from 0 to 100%. We calculated iNMBs by multiplying the incremental effects of the MRI strategy over the transrectal ultrasound guided biopsy (TRUS-GB) strategy with the Dutch willingness to pay (WTP) per quality adjusted life year (€80,000)[37] and subtracting the incremental costs. A negative iNMB, presented in red, indicates that the MRI strategy is not cost-effective, while a positive iNMB (green) indicates that the MRI strategy is the most cost-effective strategy. The figure shows that regardless of the probability of a significant tumour or the probability of a tumour, the iNMB is positive and the MRI strategy is cost-effective. For values of the sensitivity of MRGB up to and including 10%, the iNMB is negative (MRI strategy not cost-effective), while for values of 20% and higher the iNMB is positive, implying that the MRI strategy is cost-effective. Abbreviations: MRGB = MR guided biopsy; MRI = magnetic resonance imaging; TRUS = transrectal ultrasound.

DISCUSSION

The results of our model suggest that the MRI strategy is cost-effective in diagnosing PCa compared with the TRUS-GB strategy, assuming a sensitivity of MRGB of 20% or higher. Although the MRI strategy is initially more expensive, these extra costs are compensated for by reducing treatment costs resulting from fewer false positives and a better estimation of the tumour aggressiveness. The improvement in QALYs is achieved by preventing unnecessary radical treatment of insignificant tumours (with a reduced QoL without improved survival) and decreasing the chance of detecting significant tumours late (with reduced survival). The MRI strategy has the highest probability to be cost-effective at WTP values higher than €1000.

To our knowledge, this is the first paper comparing the cost-effectiveness of the MRI strategy with the TRUS-GB strategy. Recently, the Aberdeen Health Technology Assessment Group[38] published an extensive report on the diagnostic accuracy and cost-effectiveness of MR spectroscopy and other mp-MRI techniques to direct TRUS-GB compared with systematic extended-cores TRUS-GB. Their results show that a strategy using MRI may be cost-effective compared with systematic TRUS-GB. Although the conclusions are in line with this paper, the focus in their report is on a different population: patients with prior negative TRUS-GB instead of the biopsy-naive population in our paper. Furthermore, (mp-) MRI is used as a technique to aid targeted TRUS-GB instead of using MRGB.

The major strength of our model is that we used input data from a literature review regarding the effectiveness of systematic TRUS-GB and a meta-analysis of mp-MRI. Furthermore, we used costs and prevalence data from daily practice. Uncertain input parameters were varied in sensitivity and threshold analyses to show the impact of changes in these parameters on the results.

Some potential limitations should also be discussed. The sensitivity of MRGB was estimated to be 90% in the base case analysis and varied in the threshold analysis. We consider this a realistic estimate, because MRGB is performed under image guidance, with verification of needle position coming from a confirmation scan with the needle in place. Threshold analysis reveals that sensitivity of MRGB needs to be above 20% for the MRI strategy to be cost-effective.

We did not take into account the potential of mp-MRI for enabling the more favourable focal MR-guided treatments because these are not standardized in daily practice yet. Therefore, we assumed a pathway with similar treatment options for both the MRI and TRUS-GB strategies. Moreover, treatment behaviour may change in the future, namely, insignificant tumours will probably be followed periodically instead of treated radically more and more.

Costs of direct treatment or biopsy complications (eg, impotence and incontinence) were not included in our model. The estimated cost-effectiveness is therefore a conservative estimation. It can be hypothesized that cost-effectiveness becomes more in favour of the MRI strategy when these costs are considered in the analysis, because the MRI strategy potentially enables better stratification and reduction of over-treatment and uses fewer biopsy needles. Although the hospitalization rate after biopsy is reported to be comparable

in both groups, with 0-6.3 % using TRUS-GB[6] and 2% using MRGB[12], MRGB uses fewer needles, and reduces unnecessary biopsies.[16] Another assumption was that a negative MRGB or TRUS-GB was not followed by further biopsies, whereas in clinical practice, TRUS-GB is commonly repeated when clinical suspicion remains. Limited literature suggests that repeat MRGB might be an unlikely clinical occurrence, while a repeat MRI could be part of a follow-up strategy.[16]

In the present study, we were interested in the cost-effectiveness of a combination of two techniques (mp-MRI and MRGB) compared with the standard of care (TRUS-GB). Other biopsy techniques could also follow mp-MRI in improving PCa diagnosis and management, ie, targeted cognitive or ultrasound-MRI fusion techniques or a targeted transperineal approach. Although it is beyond the scope of the present study, the model can also be used for other targeted biopsy techniques. If these techniques have a similar accuracy, they would become cost-effective compared with the TRUS-GB strategy. The question then remains, which targeted biopsy technique shows the highest accuracy and which is most cost-effective.

We used Dutch cost data to estimate the cost-effectiveness, so the results may not be applicable in other countries. Given the detailed presentation of the model and its input parameters, those interested can assess the transferability of the results to their specific situation.[39] Furthermore, appropriate allocation of health care resources is becoming increasingly complex as a result of increasing patient life expectancy and increasing health care costs per patient coupled with diminishing resources caused by the global financial crisis. In this setting, and because there is no universally accepted WTP threshold among countries, we analysed the results over a broad range of WTP values.

This model is, nonetheless, a valuable instrument for assessing the uncertainties and gaps in knowledge in the currently available literature and can serve as reference point for assessing the studies that should be conducted to improve this model-based approach. When a new technique is considered cost-effective, the next step is to evaluate the feasibility of implementing this new technology in daily clinical practice. More research, preferably including a direct prospective comparison of the accuracy of the MRI and TRUS-GB strategies, and research on costs and QALYs of newly developed (focal) treatment options, and targeted biopsy techniques are needed to further validate the accuracy and reliability of our model. Furthermore, although our model reflects the average population of men who have a suspicion of PCa, future analyses should take important prognostic factors such as age and comorbidity into account.

CONCLUSIONS

Our results suggest that the MRI strategy is cost-effective compared with the standard of care using TRUS-GB, despite uncertainty around the presented cost-effectiveness estimates. The total costs of the MRI strategy are almost equal with standard of care, while potential reduction of over-diagnosis and over-treatment with the MRI strategy leads to an improvement in the QoL of PCa patients.

ACKNOWLEDGEMENTS

The authors acknowledge Jeroen Bosch, Lily van den Broeke, Wim de Geus, Harold Kerstens, and Iman Osman for providing input for our model. We thank Jonathan Richenberg for his critical review during the drafting of our manuscript.

REFERENCES

1. Elabbady AA, Khedr MM. Extended 12-core prostate biopsy increases both the detection of prostate cancer and the accuracy of Gleason score. *European urology* 2006; 49(1): 49-53; discussion.
2. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *Journal of the National Cancer Institute* 2009; 101(6): 374-83.
3. Welch HG, Black WC. Overdiagnosis in cancer. *Journal of the National Cancer Institute* 2010; 102(9): 605-13.
4. Korfage IJ, Essink-Bot ML, Borsboom GJ, et al. Five-year follow-up of health-related quality of life after primary treatment of localized prostate cancer. *International journal of cancer Journal international du cancer* 2005; 116(2): 291-6.
5. Ukimura O, Coleman JA, de la Taille A, et al. Contemporary role of systematic prostate biopsies: indications, techniques, and implications for patient care. *European urology* 2013; 63(2): 214-30.
6. Loeb S, Vellekoop A, Ahmed HU, et al. Systematic Review of Complications of Prostate Biopsy. *European urology* 2013.
7. Heijnsdijk EA, der Kinderen A, Wever EM, Draisma G, Roobol MJ, de Koning HJ. Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer. *British journal of cancer* 2009; 101(11): 1833-8.
8. Heijnsdijk EA, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. *The New England journal of medicine* 2012; 367(7): 595-605.
9. Moyer VA, Force USPST. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine* 2012; 157(2): 120-34.
10. Selnaes KM, Heerschap A, Jensen LR, et al. Peripheral zone prostate cancer localization by multiparametric magnetic resonance at 3 T: unbiased cancer identification by matching to histopathology. *Investigative radiology* 2012; 47(11): 624-33.
11. Turkbey B, Pinto PA, Mani H, et al. Prostate cancer: value of multiparametric MR imaging at 3 T for detection--histopathologic correlation. *Radiology* 2010; 255(1): 89-99.
12. Hoeks CM, Schouten MG, Bomers JG, et al. Three-Tesla magnetic resonance-guided prostate biopsy in men with increased prostate-specific antigen and repeated, negative, random, systematic, transrectal ultrasound biopsies: detection of clinically significant prostate cancers. *European urology* 2012; 62(5): 902-9.
13. Hambrock T, Hoeks C, Hulsbergen-van de Kaa C, et al. Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort. *European urology* 2012; 61(1): 177-84.
14. Siddiqui MM, Rais-Bahrami S, Truong H, et al. Magnetic Resonance Imaging/Ultrasound-Fusion Biopsy Significantly Upgrades Prostate Cancer Versus Systematic 12-core Transrectal Ultrasound Biopsy. *European urology* 2013.
15. Ahmed HU, Kirkham A, Arya M, et al. Is it time to consider a role for MRI before prostate biopsy? *Nature reviews Clinical oncology* 2009; 6(4): 197-206.
16. Moore CM, Robertson NL, Arsanious N, et al. Image-guided prostate biopsy using magnetic resonance imaging-derived targets: a systematic review. *European urology* 2013; 63(1): 125-40.
17. Hunink MG, Krestin GP. Study design for concurrent development, assessment, and implementation of new diagnostic imaging technology. *Radiology* 2002; 222(3): 604-14.
18. Heidenreich A, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. *European urology* 2011; 59(1): 61-71.
19. Thompson I, Thrasher JB, Aus G, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. *The Journal of urology* 2007; 177(6): 2106-31.

20. Guidelines for pharmaco-economic research (in Dutch). Diemen: Health Care Insurance Board, 2006.
21. Kvale R, Moller B, Wahlqvist R, et al. Concordance between Gleason scores of needle biopsies and radical prostatectomy specimens: a population-based study. *BJU international* 2009; 103(12): 1647-54.
22. Kranse R, Roobol M, Schroder FH. A graphical device to represent the outcomes of a logistic regression analysis. *The Prostate* 2008; 68(15): 1674-80.
23. Hambrock T, Somford DM, Hoeks C, et al. Magnetic resonance imaging guided prostate biopsy in men with repeat negative biopsies and increased prostate specific antigen. *The Journal of urology* 2010; 183(2): 520-7.
24. Franiel T, Stephan C, Erbersdobler A, et al. Areas suspicious for prostate cancer: MR-guided biopsy in patients with at least one transrectal US-guided biopsy with a negative finding--multiparametric MR imaging for detection and biopsy planning. *Radiology* 2011; 259(1): 162-72.
25. Roethke M, Anastasiadis AG, Lichy M, et al. MRI-guided prostate biopsy detects clinically significant cancer: analysis of a cohort of 100 patients after previous negative TRUS biopsy. *World journal of urology* 2012; 30(2): 213-8.
26. Taira AV, Merrick GS, Bennett A, et al. Transperineal Template-guided Mapping Biopsy as a Staging Procedure to Select Patients Best Suited for Active Surveillance. *Am J Clin Oncol* 2013; 36(2): 116-20.
27. de Rooij M, Hamoen EHJ, Futterer JJ, Barentsz JO, Rovers MM. Accuracy of multiparametric magnetic resonance imaging for prostate cancer: a diagnostic meta-analysis. *AJR American journal of roentgenology* Accepted for publication 2013.
28. Tewari A, Gamito EJ, Crawford ED, Menon M. Biochemical recurrence and survival prediction models for the management of clinically localized prostate cancer. *Clinical prostate cancer* 2004; 2(4): 220-7.
29. Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010; 28(7): 1117-23.
30. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding voor kostenonderzoek, methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg: College voor Zorgverzekeringen, geactualiseerde versie 2010.
31. Thompson IM, Jr., Goodman PJ, Tangen CM, et al. Long-term survival of participants in the prostate cancer prevention trial. *The New England journal of medicine* 2013; 369(7): 603-10.
32. Boorjian SA, Karnes RJ, Viterbo R, et al. Long-term survival after radical prostatectomy versus external-beam radiotherapy for patients with high-risk prostate cancer. *Cancer* 2011; 117(13): 2883-91.
33. Koscuiszka M, Hatcher D, Christos PJ, et al. Impact of race on survival in patients with clinically nonmetastatic prostate cancer who deferred primary treatment. *Cancer* 2012; 118(12): 3145-52.
34. Stewart ST, Lenert L, Bhatnagar V, Kaplan RM. Utilities for prostate cancer health states in men aged 60 and older. *Medical care* 2005; 43(4): 347-55.
35. Briggs AH, Weinstein MC, Fenwick EA, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2012; 15(6): 835-42.
36. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health economics* 1994; 3(5): 309-19.
37. Sensible and sustainable care (in Dutch). Zoetermeer: Council for Public Health and Health Care, 2006.
38. 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 technology assessment* 2013; 17(20): 1-281.
39. Welte R, Feenstra T, Jager H, Leidl R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *PharmacoEconomics* 2004; 22(13): 857-76.

6

Economic burden of urinary incontinence after prostate cancer treatment

Submitted as

de Rooij M, Nieuwboer W, Smit RNG, Witjes JA, Barentsz JO, Rovers MM

Economic burden of urinary incontinence after prostate cancer treatment

ABSTRACT

Introduction

Because current treatment options for localized prostate cancer (PCa) show high survival rates, post-treatment functional outcomes become increasingly important. Our aim was to provide insight into rates of and cost associated with urinary incontinence (UI) after PCa treatment.

Materials and Methods

We used anonymous and de-identified data from a Health Care Insurer to identify a cohort of men who had a suspicion of PCa. The database provided unique and detailed information on rates and reimbursements of UI.

Results

We identified 2834 men who underwent treatment or follow-up for PCa. UI rates in the first year varied from 8.8% for the conservative management strategy to 80.4% for men who underwent laparoscopic prostatectomy (mean for all management strategies 22.6%). Costs per patient ranged from €112 for men who underwent radiotherapy to €283 for laparoscopic prostatectomy (mean costs per patient with UI €210). In the second year after treatment or follow-up, UI rates varied from 9.2% for the conservative management strategy to 40.0% for men who underwent laparoscopic prostatectomy (mean 14.6%). Costs varied from €164 per patient who underwent conservative treatment to €292 for laparoscopic prostatectomy (mean €219).

Conclusions

The high rates and costs of UI show the extend of the burden, which is likely to increase in the future because of improved survival and earlier detection.

INTRODUCTION

In Europe, prostate cancer (PCa) has emerged as the most common non-dermatological cancer among men, with still rising incidence rates.[1] Because current treatments options for localized PCa show high survival rates, decision making is increasingly influenced by post-treatment quality of life (QoL) and adverse effects.[2-4] Urinary incontinence (UI) is a common adverse effect of PCa treatment.

Reported proportions of men who experience post-treatment UI vary amongst different treatment options, pre-treatment characteristics, and definitions that are being used.[5,6] Reliable assessment of UI is important, because the condition is often irreversible when it is still present after several years.[5] This will have implications for the QoL of men after treatment, but will also have major economic consequences. Last decades, the health care expenditures for PCa dramatically increased due to improved survival and longer life expectancy. Also, the increase of prostate specific antigen (PSA) driven detection of PCa, which can lead to over-diagnosis and over-treatment of PCa, contributes to the rising economic burden.[7,8] So far, UI rates after PCa treatment are studied in follow-up studies, but a population based assessment of UI rates and accompanied societal costs is still lacking. We had the unique possibility to assess UI rates and costs from a health insurance database with representative information of approximately 17% of the Dutch population.[9,10]

MATERIALS AND METHODS

The health insurance database of Achmea (AHD) comprises detailed individual participant information on diagnosis/treatment codes (DBCs) for reimbursement of hospital care, but also registration of pharmacy, and costs of other health care deliverables. Health insurance in the Netherlands is mandatory, and health care insurers have to offer a universal package for everyone, regardless of age or state of health. In contrast to many other European systems, the Dutch government is responsible for the accessibility and quality of the health care system in the Netherlands. Therefore, medical care is more or less similar for all inhabitants, and management of PCa patients is not controlled by the insurance company. Within the database, we selected men who underwent a PSA test in 2007. All men with a PSA test or a DBC registration code for PCa in 2006 were excluded to select a population of men at the beginning of their PCa management strategy.

Figure 1 shows the flow chart of the identification of different management strategies for men that underwent a PSA test. Men who were not insured during the complete inclusion period (2007 through 2011), and men under the age of 30 were also excluded.

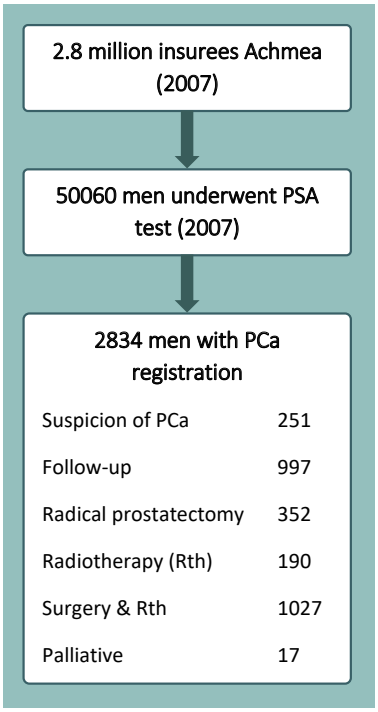


Figure 1 Flow chart of the identification of management strategies of men who underwent a PSA test. Abbreviations: PCa = prostate cancer; PSA = prostate specific antigen; Rth = radiotherapy.

Six different management strategies were identified by their corresponding DBC codes: (1) men with a suspicion of PCa; these men had only one PCa DBC registration code for conservative treatment (outpatient only); (2) follow-up (active surveillance/watchful waiting); men with more than one PCa DBC registration code for conservative treatment were assumed to be included in a follow-up management strategy; (3) radical prostatectomy (open/laparoscopic); men who had a DBC registration code with a clinical episode, specified as surgical (open/laparoscopic); (4) radiotherapy; similarly but with a registration code for radiotherapy (technical details not specified); (5) prostatectomy and radiotherapy (open/laparoscopic); codes for both surgery and radiotherapy; (6) palliative treatment; men with registration codes specific for PCa treatment at the department of internal medicine, suggestive for palliative management strategy. Data of all included men were analysed to assess the proportion of patients with treatment related UI and the associated health

insurance reimbursements. The reimbursements included urinary incontinence materials like diapers, catheters, and specific physiotherapy programs for urinary incontinence. Costs for diagnostic procedures to investigate incontinence and surgical interventions for urinary incontinence were not included. To ensure that we only analysed treatment related UI, men with reimbursements for UI before treatment or follow-up were not included. For all management strategies, the influence of transurethral resection of the prostate in the history was assessed.

Ethical approval was not required, as we only used anonymous and de-identified data from the AHD.

RESULTS

In total, 50060 men with a PSA test in 2007, but without a previous PSA test or DBC of PCa, were analysed. A total of 2834 of these men underwent one of the six identified PCa management strategies.

Table 1 shows the identified management strategy with associated UI rates and costs for the first and second year after treatment or follow-up. UI was most common after prostatectomy; 64.5% and 80.4% had UI reimbursements in the first year post-treatment for open and laparoscopic prostatectomy, respectively. These UI rates decreased to 29.9% and 40.0% in the second year after treatment. UI rates were 13.4% in the first year after radiotherapy, and 9.7% in the second year. In the first year after palliative treatment the UI rate was 29.4%, and 17.7% in the second year. For patients with both prostatectomy and radiotherapy UI was present in 38.8% and 59.8% of men with an open and laparoscopic approach, respectively. After two years, these pathways showed rates of 37.9% and 29.9%, respectively. Patients with a suspicion of prostate cancer showed an UI rate of 20.3% after the first year of follow-up, which decreased to 10.0% in the second year. The group with conservative treatment showed the lowest UI rates: 8.8% and 9.2% after the first and second year of follow-up.

Table 1 Post-treatment urinary incontinence rates and costs

		First year of follow-up/post-treatment				Second year of follow-up/post-treatment			
Management strategy	Group size (n)	Urinary incontinence (n (%))	Urinary incontinence costs per patient with UI in € (mean)	Total urinary incontinence costs for Dutch population in € (mean)*	Costs per 10,000 men in €	Urinary incontinence (n, (%))	Urinary incontinence costs per patient with UI in € (mean)	Total urinary incontinence costs for Dutch population in € (mean)*	Costs per 10,000 men in €
Clinical suspicion	251	51 (20.3)	157.12	46812	57.87	25 (10.0)	176.09	25719	31.79
Conservative treatment	997	88 (8.8)	155.86	80130	99.06	92 (9.2)	163.88	88082	108.89
Prostatectomy									
Open	107	69 (64.5)	258.25	104101	128.70	32 (29.9)	255.64	47791	59.08
Laparoscopic	245	197 (80.4)	282.93	325627	402.56	98 (40.0)	292.17	167279	206.80
Prostatectomy + Radiotherapy									
Open	103	40 (38.8)	240.08	56103	69.36	39 (37.9)	241.86	55106	68.13
Laparoscopic	87	52 (59.8)	244.00	74126	91.64	26 (29.9)	269.40	40921	50.59
Radiotherapy	1027	138 (13.4)	111.64	90006	111.27	100 (9.7)	177.03	103425	127.86
Palliative	17	5 (29.4)	255.84	7473	9.24	3 (17.7)	211.62	3709	4.59
Total	2834	640 (22.6)	209.78	784379	969.69	415 (14.6)	219.44	532031	657.72

* In 2007 the Achmea Health Database comprised information of 2.8 million insurants. The Dutch population was 16.36 million (8.09 million men) in 2007. We used an extrapolation factor of 16.36/2.8 = 5.84. UI = urinary incontinence

Table 1 displays health insurance UI costs of the identified management strategy as yearly UI costs per patient, and UI costs per management strategy, extrapolated to the total Dutch population (16.36 million in 2007[10]), and per 10000 men. Yearly, UI reimbursement costs per patient ranged from €112 (radiotherapy group) to €283 (laparoscopic prostatectomy group). Total costs per management strategy extrapolated to costs per 10000 men ranged from €9 in men who underwent palliative treatment to €402 in the laparoscopic prostatectomy group. These sum up to a total UI cost of €784379 for the total Dutch population in the first year (€967 per 10000 men). The second year, the total treatment related UI costs extrapolated to the Dutch population were €532031 (€658 per 10000 men). In both years, the prostatectomy groups were responsible for the major part of the UI costs.

DISCUSSION

In this research report, we highlight the rates and economic burden of UI related to PCa treatment. We used longitudinal health insurance data of the Achmea Health Database (AHD), which offered us the unique possibility to extract management strategies and UI rates with corresponding UI reimbursement costs.

Among the different treatment options UI rates were highest after prostatectomy, which is in accordance with current literature.[6,11] The higher UI rates for laparoscopic prostatectomy as compared to an open procedure, especially in the first year, could possibly be explained by difference in experience with these techniques.

An unexpected finding was the lower UI rate in the first year for men who underwent both prostatectomy and radiotherapy, which we cannot explain by the current data. Another unexpected finding was the de novo UI in men without surgical management or radiotherapy (eg, in the group of men with suspicion of PCa or follow-up only). This could possibly be explained by a referral for both urinary incontinence and work-up for suspicion of PCa. These unexpected findings illustrate one of the restrictions of the AHD: the limited availability of clinical data, such as indications, results of diagnostic tests, and cause of death[9], which makes the identification of the treatment pathways challenging. This also resulted in the inability to provide detailed patient characteristics of the included men (for instance data on PSA level, cancer stage, and co-morbidity were missing).

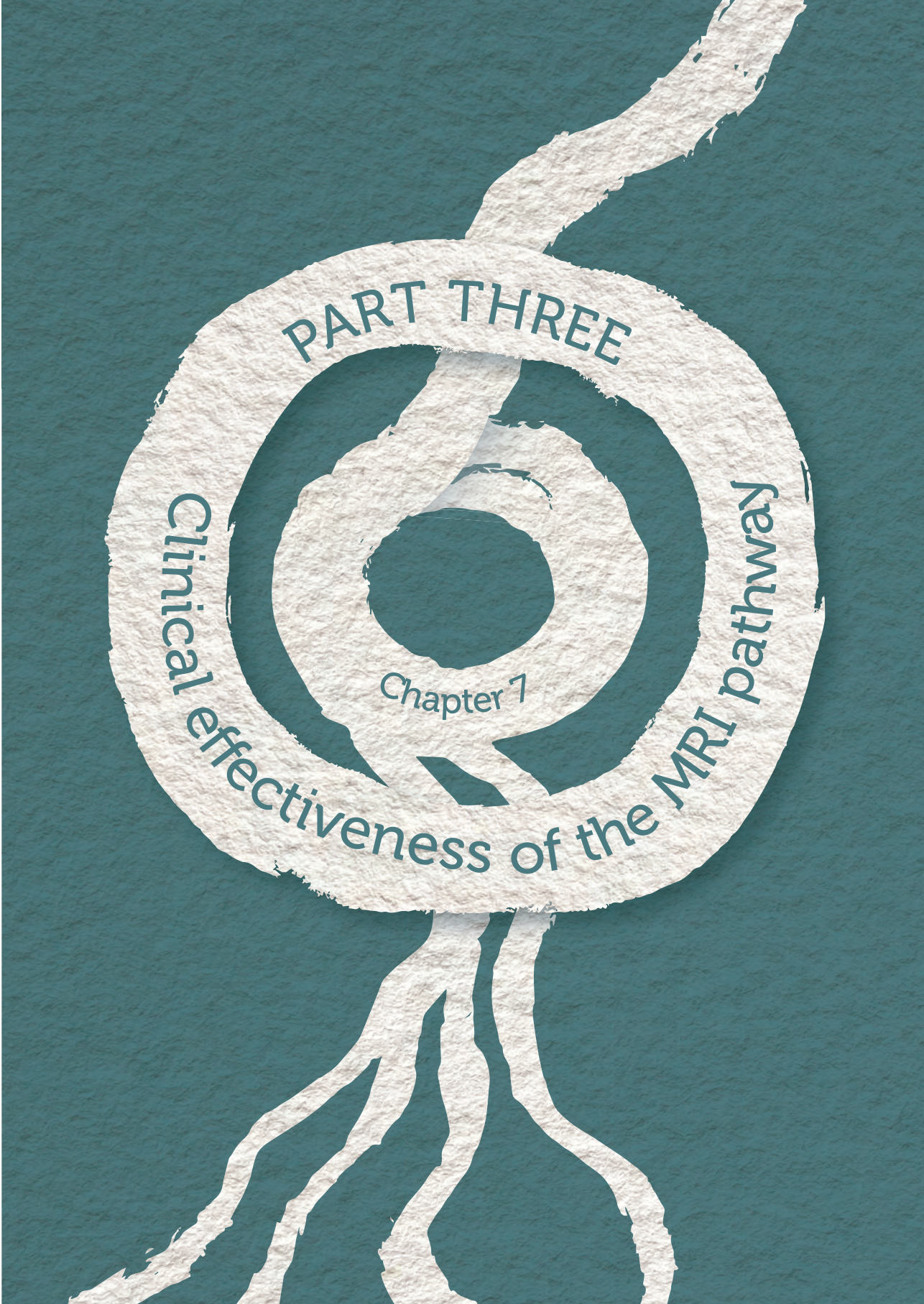
Identification of UI with this population based database shows an estimate of the magnitude of the problem, but detailed information on the severity of the UI is lacking. Costs per patient can be a proxy for the severity, but it is likely that the presented total costs of UI are an underestimation of the societal problem, because UI is still a taboo for many men. Although we present Dutch cost information, the costs per 10000 men makes extrapolation and comparability to other countries possible.

CONCLUSION

The high rates and costs of UI found in the current study, particularly after prostatectomy, show an estimation of the magnitude of the UI burden. The burden is likely to increase in the future because of improved survival and earlier detection of PCa.

REFERENCES

1. Center MM, Jemal A, Lortet-Tieulent J, et al. International variation in prostate cancer incidence and mortality rates. *European urology* 2012; 61(6): 1079-92.
2. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. *The New England journal of medicine* 2014; 370(10): 932-42.
3. Xiong T, Turner RM, Wei Y, Neal DE, Lyratzopoulos G, Higgins JP. Comparative efficacy and safety of treatments for localised prostate cancer: an application of network meta-analysis. *BMJ open* 2014; 4(5): e004285.
4. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. *The New England journal of medicine* 2012; 367(3): 203-13.
5. Resnick MJ, Koyama T, Fan KH, et al. Long-term functional outcomes after treatment for localized prostate cancer. *The New England journal of medicine* 2013; 368(5): 436-45.
6. Johansson E, Steineck G, Holmberg L, et al. Long-term quality-of-life outcomes after radical prostatectomy or watchful waiting: the Scandinavian Prostate Cancer Group-4 randomised trial. *The lancet oncology* 2011; 12(9): 891-9.
7. Roehrborn CG, Black LK. The economic burden of prostate cancer. *BJU international* 2011; 108(6): 806-13.
8. Loeb S, Bjurlin MA, Nicholson J, et al. Overdiagnosis and overtreatment of prostate cancer. *European urology* 2014; 65(6): 1046-55.
9. Smeets HM, de Wit NJ, Hoes AW. Routine health insurance data for scientific research: potential and limitations of the Agis Health Database. *Journal of clinical epidemiology* 2011; 64(4): 424-30.
10. Centraal Bureau voor de Statistiek Den Haag/Heerlen.
11. Jarosek SL, Virnig BA, Chu H, Elliott SP. Propensity-weighted Long-term Risk of Urinary Adverse Events After Prostate Cancer Surgery, Radiation, or Both. *European urology* 2014.


PART THREE
Chapter 7
Clinical effectiveness of the MRI pathway

7

Prospective study of diagnostic accuracy comparing prostate cancer detection by TRUS-GB versus MRI pathway in biopsy-naive men

Published as

Pokorny MR, de Rooij M, Duncan E, Schröder FH, Parkinson R, Barentsz JO, Thompson LS

Prospective study of diagnostic accuracy comparing prostate cancer detection by transrectal ultrasound-guided biopsy versus magnetic resonance (MR) imaging with subsequent MR-guided biopsy in men without previous prostate biopsies

European Urology. 2014 Jul;66(1):22-9

ABSTRACT

Background and objective
The current diagnosis of prostate cancer (PCa) uses transrectal ultrasound guided biopsy (TRUS-GB). TRUS-GB leads to sampling errors causing delayed diagnosis, over-detection of indolent PCa, and misclassification. Advances in multiparametric MRI (mp-MRI) suggest that imaging and selective MR-guided biopsy (MRGB) may be superior to TRUS-GB. The objective was to compare the diagnostic efficacy of the MRI pathway with TRUS-GB.

Design, setting, and participants
A total of 223 consecutive, biopsy-naive men referred to a urologist with elevated PSA participated in a single-institution, prospective, investigator-blinded, diagnostic study from June 2012 through January 2013.

Intervention
All participants had mp-MRI and TRUS-GB. Men with equivocal or suspicious lesions on mp-MRI also underwent MRGB.

Outcome measurements and statistical analysis
The primary outcome was PCa detection. Secondary outcomes were histopathologic details of biopsy and radical prostatectomy specimens, adverse events, and MRI reader performance. Sensitivity, specificity, and negative/positive predictive values (NPV/PPV) were estimated and basic statistics presented by number (percentage) or median (interquartile range).

Results and limitations
Of 223 men, 142 (63.7%) had PCa. TRUS-GB detected 126 cases of PCa in 223 men (56.5%) including 47 (37.3%) classed as low-risk. MRGB detected 99 cases of PCa in 142 men (69.7%) with equivocal or suspicious mp-MRI, of which 6 (6.1%) were low-risk. The MRGB pathway reduced the need for biopsy by 51%, decreased the diagnosis of low-risk PCa by 89.4%, and increased the detection of intermediate/high-risk PCa by 17.7%. The estimated NPVs of TRUS-GB and MRGB for intermediate/high-risk disease were 71.9% and 96.9%, respectively. The main limitation is the lack of long follow-up.

Conclusions
We found that mp-MRI/MRGB reduces the detection of low-risk PCa and reduces the number of men requiring biopsy while improving the overall rate of detection of intermediate/high-risk PCa.

Patient summary
We compared the results of standard prostate biopsies with a MRI image-based targeted biopsy diagnostic pathway in men with elevated PSA. Our results suggest patient benefits of the MRI pathway. Follow-up of negative investigations is required.

INTRODUCTION

PCa is the most common male malignancy and the second most common cause of male cancer related death.[1] Randomized trials have shown that early detection, through prostate specific antigen (PSA) screening, can alter the natural history of the disease and reduce mortality.[2] However, this benefit is associated with the diagnosis of many indolent tumours, for which radical treatment leads to an adverse impact on quality of life without altering survival.[3-7] Population-based reports suggest little disconnect between diagnosis and treatment.[8-10] The over-diagnosis and over-treatment of PCa has caused various professional organisations to review their PSA screening guidelines,[11] potentially reversing recent declines in disease-specific mortality.[12]

Another approach to minimize over-treatment, would be to reduce the over-diagnosis of low-risk PCa. Urologists use PSA followed by systematic transrectal ultrasound guided biopsy (TRUS-GB) rather than imaged based diagnosis due to poor discrimination of PCa with transrectal ultrasound (TRUS). Due to the high prevalence of low-risk PCa, the TRUS-GB diagnostic pathway finds many indolent tumours. TRUS-GB also causes difficulties in managing patients with high PSA values but benign biopsies, and it also misclassifies the volume or risk of approximately a third of cases of biopsy-detected PCa when compared with whole-mount pathology.[13-15]

Advances in 3 Tesla multi-parametric MRI (mp-MRI) have improved image-based diagnosis.[16-18] Also, targeted MR-guided biopsy (MRGB) has become an alternative approach to TRUS-GB. MRGB uses fewer cores than TRUS-GB and can be applied only in men with lesions suspicious for intermediate/high-risk PCa.[19-22] Although selective MRGB is an appealing pathway, few data support its reliability. It is unknown what proportion and what type of PCa would be missed by omitting biopsy from men with normal mp-MRI scans. With this in mind, we designed a prospective diagnostic study to compare selective MRGB and unselected TRUS-GB in men with an elevated PSA.

MATERIALS AND METHODS

Recruitment, imaging, and biopsy

In this prospective single-centre diagnostic study, 226 biopsy-naive subjects with concerning PSA levels and/or an abnormal digital rectal examination (DRE) were consecutively enrolled by referral from urologists from July 2012 through January 2013.

All subjects underwent prostate mp-MRI performed at 3 Tesla (Magnetom Skyra, Siemens) without an endorectal coil. The mp-MRI detection protocols and technique for MRGB have been reported.[23] Exclusion criteria are stated in Figure 1; three subjects were excluded.

The mp-MRI images were scored independently by three readers (1 year, 1 year, and 19 years of experience) using the validated Prostate Imaging Reporting and Data System (PI-RADS): from 1 (low) to 5 (high) according to the likelihood of significant PCa being present.[23-25] The two less experienced readers were trained for 2 weeks at a reference centre and read at least 100 cases with feedback. Disagreements were resolved by consensus.

Patients returned for prostate biopsy at a second visit. Subjects with equivocal (PI-RADS 3) and intermediate/high-risk (PI-RADS 4/5) lesions underwent MRGB followed within 30 minutes by TRUS-GB performed by a urologist blinded to the mp-MRI findings and the MRGB procedure. A 12-core TRUS-GB was performed in a standard paired sextant pattern. Any lesions seen on TRUS were targeted using the core for the relevant prostate zone. Subjects with normal mp-MRI scans (PI-RADS 1/2) underwent TRUS-GB only.

MRGB fulfilled the Standards of Reporting for MRI-targeted Biopsy Studies (START) recommendations.[26] Each mp-MRI suspicious or equivocal lesion was biopsied using two to three cores. All biopsy specimens underwent evaluation by a certified urogenital histopathologist blinded to the origin of each core.

Institutional approval was obtained, and all men provided written informed consent. Figure 1 shows study the design.

Trial outcomes

The primary outcome of this study was the number of men detected with PCa including stratification into low- and intermediate/high-risk disease.[8-10] Low-risk PCa was defined as either low-volume Gleason score 3+3 or very low volume Gleason score 3+4. A modality-specific biopsy-based definition was required because of the different number of cores obtained from TRUS-GB and MRGB. Insignificant cancer at radical prostatectomy histology was defined using active surveillance criteria (total tumour volume <0.7 ml and Gleason score ≤3+4) to be consistent with the study biopsy risk stratification schedule.[27] Secondary outcomes included histologic details of the biopsies and identified tumours, MRI reader performance, histology of radical prostatectomy, and adverse events.

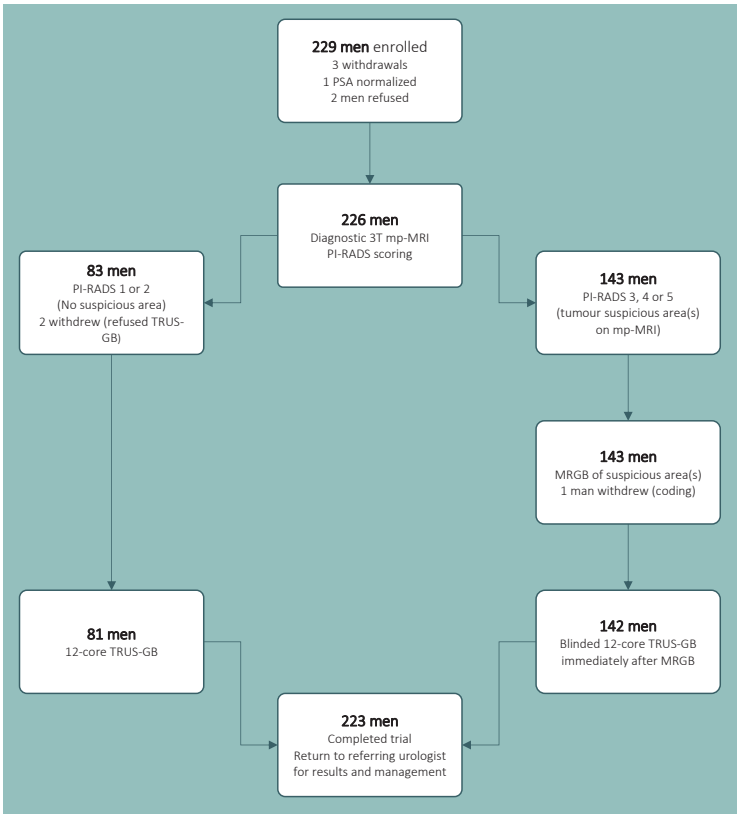


Figure 1 Flow diagram depicting the stratified detection design of the diagnostic study showing TRUS-GB pathway on the left and MRI pathway on the right. Exclusion criteria: patients with prior prostate biopsy or known existing prostate cancer, 5α-reductase inhibitor therapy, contraindication to MRI (eg, pacemaker) or allergy to gadolinium contrast. Abbreviations: mp-MRI = multi-parametric magnetic resonance imaging; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate specific antigen; TRUS-GB = transrectal ultrasound guided biopsy.

Statistical analyses

We compared mp-MRI/MRGB with TRUS-GB using sensitivity, specificity, negative and positive predictive values (NPVs, PPVs).[28,29] The maximum likelihood estimates (MLEs) of PPV and NPV are obtained from the MLEs of the sensitivity, specificity, and prevalence rate using the Baye theorem. These estimates are straightforward to calculate when the true disease state is known. However, this would require a "perfect" reference test (eg, radical prostatectomy). Because TRUS-GB is not perfect, sensitivity and specificity of the TRUS-GB must also be estimated. This is problematic because there are five variables to estimate but only 3 degrees of freedom provided by the cross-tabulation of the two tests. The solution proposed by Enøe et al.[30] is to subset the data based on some covariate, such that the sub-populations have different prevalence rates, and perform the cross-tabulations for each

sub-population. This solution uses the model introduced by Hui and Walter,[31] who derived explicit forms of the MLEs of these six parameters. This model was used to estimate test parameters for the two diagnostic pathways.

RESULTS

Patients and tumours

A total of 223 subjects were available for the final analysis (Figure 1). These were typical for a white population at risk of PCa (Table 1) with a median age of 63 years (interquartile range (IQR) 57–68), median serum of PSA 5.3ng/mL (IQR 4.1–6.6), and median prostate volume of 41ml (IQR 30–59). DRE was suspicious of PCa in 40 men (17.9%). The mp-MRI was suspicious for PCa (PI-RADS 4/5) in 109 (49%), equivocal (PI-RADS 3) in 33 (15%), and identified no abnormality (PI-RADS 1/2) in the remaining 81 men (36%).

Table 1 Summary statistics relating to subjects and MRI findings.

	n (223)	%
PSA		
Median (IQR)	5.3 (4.1-6.6)	
0-3.9	48	21.5
4-9.9	156	70.0
10-26	19	8.5
Age		
Median (IQR)	63 (57-68)	
<60	73	33
≥60	150	67
Prostate volume		
Median (IQR)	41 (30-59)	
<30 ml	50	22.4
30-50 ml	95	42.6
>50 ml	78	35.0
DRE		
Normal	183	83
Abnormal	40	17
mp-MRI score		
PI-RADS 1/2	81	36
PI-RADS 3	33	15
PI-RADS 4/5	109	49

Abbreviations: DRE = digital rectal exam; IQR = interquartile range; mp-MRI = multi-parametric magnetic resonance imaging; PSA = prostate specific antigen.

TRUS-GB

TRUS-GB was performed in 223 men and PCa detected in 126 (56.5%, Tables 2 and 3). Of these, 47 (37.3%) were low-risk; 79 (62.7%) were intermediate/high-risk. Most patients with PCa (n=80, 63.5%) had predominant Gleason 3 architecture (Gleason 3+3 or 3+4).

Of 2672 TRUS-GB cores, 401 (15%) contained PCa. In PI-RADS 4/5 scans, PCa was found in 324 of 1305 TRUSB cores (24.8%). For those with PI-RADS 1/2 scans, PCa was found in 4.5% of TRUS-GB cores.

Table 2 Analysis of cancer detection by TRUS-GB, MRGB, and combined modality histology.

Histology result	TRUS-GB	MRGB	Combined Histology
Benign	97	43	81
Low-risk cancer	47	6	34
Intermediate/high-risk cancer	79	93	108
Total	223	142	223

This table shows the histology results of both diagnostic pathways separately and shows a combined histology. The combined histology is determined from the highest cancer risk found with either TRUS-GB or MRGB. Of all tumours, 71.5% were located in the peripheral zone (base 35.7%, mid-gland 50.0%, apex 14.3%). The remainder (28.5%) were located in the anterior gland or anterior horn of the peripheral zone.
Abbreviations: MRGB = magnetic resonance image guided biopsy; TRUS-GB = transrectal ultrasound guided biopsy.

MRI pathway

MRGB was performed in 142 men and detected PCa in 99 (69.7%), of which 6 (6.1%) were low-risk and 93 (93.9%) were intermediate/high-risk. The proportion of MRGB-detected intermediate/high-risk PCa was significantly higher than TRUS-GB (62.7%, $p<0.01$).
PCa detection rates were lower in PI-RADS 3 than in PI-RADS 4/5 scans (15.1% vs. 86.2%; $p<0.001$). Although only five men with PI-RADS 3 scans had PCa, three of them had intermediate/high-grade PCa. Of 417 MRGB cores, 235 (56.4%) contained PCa. This was significantly higher than for the 2672 TRUS-GB cores of which 401 (15%) contained PCa. In men with PI-RADS 4/5 (n=109), PCa was found in 225 of 322 cores (69.9%).

Comparison of detection rates MRI pathway and TRUS-GB

We found that mp-MRI/MRGB detected 6 cases of low-risk PCa (6.1%); TRUS-GB detected 47 cases (62.7%) of low-risk PCa ($p<0.001$). MRGB detected intermediate/high-risk PCa in 29 men that were either missed (n=16) or misclassified as low-risk by TRUS-GB (n=13) (Table

4). In the 16 PCa missed by TRUS-GB, Gleason scores were 3+3 (n=3), 3+4 (n=8), 4+3 (n=3), 4+4 (n=1), and 4+5 (n=1). In the misclassified cohort, risk adjustment occurred for volume in seven patients and for both grade and tumour volume in six patients. A discordance analysis is presented in the Supplemental material which is available online. TRUS-GB detected intermediate/high-risk PCa in 15 men who were either not diagnosed on mp-MRI (n=5) or missed on MRGB (n=10). In men with normal mp-MRI, PCa was detected by TRUS-GB in 25 (30.8%), and most (n=20) had low-risk PCa. A detailed description of the tumours missed or not detected by the MRI pathway is presented in the Supplemental material, which is available online.

Restricting MRGB to PI-RADS 4/5 instead of PI-RADS 3-5 lesions increased the reduction of men requiring biopsy from 36.3% to 51.1%, increased the reduction of the number of men diagnosed with low-risk PCa from 87.2% to 89.4%, and decreased the yield of intermediate/high-risk PCa from 17.7% to 12.6%. Finally, the reduction in biopsy needle cores analysed increased from 84.4% to 87.9% (Table 4). Table 5 shows details of the performance characteristics of each pathway. We found that the mp-MRI/MRGB outperforms TRUS-GB except for the PPV, which differs by less than 1%. Sensitivity and NPV is much smaller for the TRUS-GB than it is for the mp-MRI/MRGB, implying that TRUS-GB is more prone to diagnosing patients as having no intermediate/high-risk PCa when intermediate/high-risk PCa is present. The lower specificity indicates that TRUS-GB is more likely to diagnose patients as having low-risk PCa, when in fact intermediate/high-grade PCa is present.

Table 3 Detailed biopsy core analysis.

	TRUS-GB								MRGB					
	Total		PI-RADS 1/2		PI-RADS 3		PI-RADS 4/5		Total		PI-RADS 3		PI-RADS 4/5	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Total	223		81		33		109		142		33		109	
PSA														
Median	5.3		47		4.7		5.9		5.7		4.7		5.9	
IQR	2.5		2.1		2.5		2.9		2.8		2.5		2.9	
Prostate cancer														
Not detected	97	43	56	69	18	55	23	21	43	30	28	85	15	14
Low-risk	47	21	20	25	10	30	17	16	6	4	2	6	4	4
Intermediate/high-risk	79	35	5	6	5	15	69	63	93	65	3	9	90	83
Differentiation, maximum Gleason score														
3+3	38	17	16	20	10	30	12	11	15	11	3	9	12	11
3+4	42	19	6	7	5	15	31	28	42	30	2	6	40	37
4+3	26	12	3	4	0	0	23	21	21	15	0	0	21	19
≥4+4	20	9	0	0	0	0	20	18	21	15	0	0	21	19
Biopsy cores														
Total cores sampled	2672		973		394		1305		417		95		322	
Total positive cores	401	15	44	4.5	33	8.4	324	24.8	235	56.4	10	10.5	225	69.9
Gleason Sum														
6	149	37.2	30	68.2	27	81.8	92	28.4	48	20.4	6	60	42	18.7
7	193	48.1	14	31.8	6	18.2	173	53.4	142	60.4	4	40	138	61.3
8	14	3.5					14	4.3	9	3.8			9	4.0
9	41	10.2					41	12.7	36	15.3			36	16.0
10	4	1.0					4	1.2						
Maximum length of PCa														
Mean	9.02		4.75		6.73		12.87		10.49		2.55		12.87	
SD	8.23		7.52		7.75		7.05		7.47		6.14		6.07	
Cumulative length of PCa														
Mean	25.99		8.21		14.00		42.90		21.63		4.52		26.69	
SD	33.01		18.83		23.71		35.42		17.62		11.65		15.79	
Clinical stage														
Normal	183	82	74	91	28	85	81	74	109	77	28	85	81	74
cT2A	35	16	7	9	5	15	22	20	28	20	5	15	22	20
cT2B	1	0	0	0	0	0	1	1	1	1	0	0	1	1
cT2C	1	0	0	0	0	0	1	1	1	1	0	0	1	1
≥cT3	3	1	0	0	0	0	3	3	3	2	0	0	3	3

Abbreviations: IQR = interquartile range; MRGB = magnetic resonance guided biopsy; PCa = prostate cancer; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate specific antigen; SD = standard deviation; TRUS-GB = transrectal ultrasound-guided biopsy.

Table 4 Comparison of biopsy outcomes with two different approaches: (a) Subjects with PI-RADS 1/2 versus 3-5; (b) subjects with PI-RADS 1-3 versus 4/5.

a.			
	TRUS-GB pathway	MRGB pathway	Difference, %
Biopsy, no. of men	223	142	-36.3
Biopsy cores, no.	2672	417	-84.4
Low-risk PCa	47	6	-87.2
Intermediate/high-risk PCa	79	93	17.7
b.			
	TRUS-GB pathway	MRGB pathway	Difference, %
Biopsy, no. of men	223	109	-51.1
Biopsy cores, no.	2672	322	-87.9
Low-risk PCa	47	5	-89.4
Intermediate/high-risk PCa	79	89	12.6

Abbreviations: MRGB = magnetic resonance image guided biopsy; PCa = prostate cancer; TRUS-GB = transrectal ultrasound-guided biopsy.

Comparison of tumour features MRI pathway and TRUS-GB

Although avoidance of biopsy was a primary measure in this study, an important metric is the accurate determination of cancer phenotype. The proportion of cores containing cancer was 401 of 2672 (15%) for TRUS-GB compared with 235 of 417 (56.4%) for MRGB ($p<0.001$; pooled z-test). For patients who had PI-RADS 4/5 scans, these values were 24.8% for TRUS-GB and 69.9% for MRGB ($p<0.001$).

Average percent cancer core length was 32.9% for TRUS-GB and 60.6% for MRGB ($p<0.001$; two-sample t-test). In PI-RADS 4/5 scans, average percent cancer core length was 40.4% for TRUS-GB compared with 59.7% for MRGB ($p<0.001$). Supplemental Table 3 shows the number of PI-RADS 4/5 lesions per patient identified on mp-MRI (available online). Gleason 7 cancer was identified in 7.2% of all TRUS-GB cores compared with 34.1% of all MRGB cores ($p<0.001$). Gleason 7-10 cancer was identified in 9.4% of TRUS-GB cores compared with 44.8% of MRGB cores ($p<0.001$). The relative risk for identification of intermediate/high-risk cancer in men with PI-RADS 4/5 scans compared with PI-RADS 1/2 scans was 14.42 (range: 6.14-33.86).

MRI reader performance

Concordance analysis was performed to compare each MRI reader's PI-RADS score to final pathology. The results (Supplemental Table 4, available online) indicated reader

performance was equivalent. Area under the receiver operating curve was as follows: radiologist 0.85 (95% confidence interval (CI) 0.80-0.90), urologist 0.88 (95% CI 0.83-0.92), and expert radiologist 0.85 (95% CI 0.80-0.90). Accuracy did not increase during the trial, indicating that the learning curve for the readers had been completed prior to trial commencement.

Table 5 Statistical performance characteristics of transrectal ultrasound guided biopsy and magnetic resonance guided biopsy.

Parameter	Estimate	95% confidence interval
Sensitivity: (MRGB)	0.9234	(0.878, 0.953)
Sensitivity: (TRUS-GB)	0.7044	(0.639, 0.762)
Specificity: (MRGB)	0.9691	(0.934, 0.986)
Specificity: (TRUS-GB)	0.9364	(0.894, 0.963)
PPV: (MRGB)	0.9232	(0.878, 0.953)
PPV: (TRUS-GB)	0.9321	(0.889, 0.960)
NPV: (MRGB)	0.9691	(0.934, 0.987)
NPV: (TRUS-GB)	0.719	(0.654, 0.776)
LR(+): MRGB	12.908	
LR(+): TRUS-GB	5.4	
LR(-): MRGB	0.126	
LR(-): TRUS-GB	0.293	

Abbreviations: LR = likelihood ratio; MRGB = magnetic resonance image guided biopsy; NPV = negative predictive value; PPV = positive predictive value; TRUS-GB = transrectal ultrasound-guided biopsy.

Radical prostatectomy histology

Following diagnosis, 75 of the 142 men with PCa underwent radical prostatectomy. An analysis was done relating mp-MRI/MRGB and TRUS-GB diagnostic pathways to the presence of significant or insignificant cancer in the histology from these cases. Because of treatment selection bias and low numbers, concordance data are presented in Supplemental Table 5 (available online). With the mp-MRI/MRGB pathway, four men had insignificant tumours treated, 11 men had significant PCa missed, and 60 men were correctly assessed. With the TRUS-GB pathway, four men had insignificant tumours treated, 18 men had significant PCa missed, and 53 men were correctly assessed.

Adverse events

The referring urologists reported any complication promptly, without a limit on the time frame. Two subjects (0.9%; both PI-RADS 1,2) developed urosepsis after TRUS-GB. Both

had benign biopsies and recovered uneventfully. One subject required admission for haematuria after TRUS-GB. One subject experienced a vasovagal episode after MRGB.

DISCUSSION

The results of this diagnostic study support apparent patient benefits. Most importantly, when using the mp-MRI/MRGB pathway instead of TRUS-GB, the number of men diagnosed with low-risk PCa will be reduced, and at the same time, the number of men diagnosed with intermediate/high-risk PCa will be increased. Secondly, because the data indicate minimal benefit from MRGB of PI-RADS 3 lesions, the number of men who need a biopsy (ie, in PI-RADS 4-5 only) can be halved, requiring only 2-3 needles instead of the standard 10-14 needles.

Recent studies support the superiority of MR-directed biopsy over TRUS-GB to decrease detection of low-risk and increase detection of intermediate/high-risk PCa.[32-35] The studies use indirect MR-guidance techniques like cognitive or MRI/TRUS-fusion targeting that are prone to motion, segmentation, and registration errors.[36] Furthermore, these studies are limited by the use of matched cohorts instead of one single cohort or do not study a full biopsy-naive cohort, which are both strong points of our diagnostic study.

A limitation of our study is the lack of oncologic follow-up data. Reviewing the available radical prostatectomy data of our study (ie, 75 patients) should be done with caution because of treatment selection bias and low numbers. Nevertheless, more men were correctly assessed with the mp-MRI/MRGB pathway than with the TRUS-GB pathway.

Another limitation is the order of the two biopsy sessions (first MRGB, second TRUS-GB). This order makes it possible for the urologist to identify the MRGB track and take more samples from this area. Although there is no other way to do this practically without inconveniencing the patient, it could lead to a bias in favour of TRUS-GB.

Our study reveals failures of the mp-MRI/MRGB pathway. TRUS-GB detected intermediate/high-risk PCa in 15 men that were either not diagnosed on mp-MRI (n=5, PI-RADS 1,2) or missed on MRGB (n=10, PI-RADS 3-5). However, using the TRUS-GB pathway, overall more intermediate/high-risk cases of PCa were missed. Combining mp-MRI/MRGB and TRUS-GB would increase the detection rate of PCa, but the combined pathway negates the major

advantage of the MRI pathway (ie, the near elimination of the diagnosis of low-risk PCa), as well as incurring additional costs.

The MRI reader performance results (Supplemental Table 3, available online) indicate that an individual reader gets the diagnosis correct in about 85% of cases, with no significant difference between the expert radiologist, the radiologist, and the urologist. This study was performed in a well-resourced Australian hospital 18 months after introducing an mp-MRI service. Thus, with radiology support, interested urologists can learn MRI interpretation and MRGB.

Costs of new (imaging) techniques such as mp-MRI and MRGB are subject to debate. However, when using a decision analytic model to assess the cost-effectiveness of the mp-MRI/MRGB pathway compared with TRUS-GB, the MRI pathway is shown to be a cost-effective strategy when sensitivity of MRGB is higher than 20%.[37] Although upfront diagnostic costs may be higher in some countries, the ability of mp-MRI to prevent unnecessary biopsies and reduce over-treatment can lead to lower costs and higher quality of life when longer follow-up and treatment are taken into account.

It can be argued epidemiologically that the benefits of reducing over-diagnosis outweigh delaying diagnosis of a few intermediate/high-risk PCa missed at initial diagnosis, as long as a follow-up protocol is in place. Using mp-MRI/MRGB may allow the benefits of screening without adverse consequences of over-diagnosis and unnecessary treatments. Screening strategies offer men with increased PSA a 20-30% reduction of their PCa mortality[38] and significant (at least 30%) reduction of metastatic disease.[39]

Future studies with longer oncologic follow-up and comparison of the different targeted biopsy techniques are needed to assess which technique is preferable, also in terms of implementation and costs.

CONCLUSIONS

For asymptomatic men with elevated PSA, mp-MRI followed by selective use of MRGB compared with TRUS-GB reduces the detection of low-risk PCa, and it reduces the need for biopsy while improving the overall detection of intermediate/high-risk PCa.

REFERENCES

1. Center MM, Jemal A, Lortet-Tieulent J, et al. International variation in prostate cancer incidence and mortality rates. *European urology* 2012; 61(6): 1079-92.
2. Roobol MJ, Kranse R, Bangma CH, et al. Screening for prostate cancer: results of the rotterdam section of the European randomized study of screening for prostate cancer. *European urology* 2013; 64(4): 530-9.
3. Heijnsdijk EA, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. *The New England journal of medicine* 2012; 367(7): 595-605.
4. Holmstrom B, Johansson M, Bergh A, Stenman UH, Hallmans G, Stattin P. Prostate specific antigen for early detection of prostate cancer: longitudinal study. *Bmj* 2009; 339: b3537.
5. Resnick MJ, Koyama T, Fan KH, et al. Long-term functional outcomes after treatment for localized prostate cancer. *The New England journal of medicine* 2013; 368(5): 436-45.
6. Thompson IM, Ankerst DP, Chi C, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. *JAMA : the journal of the American Medical Association* 2005; 294(1): 66-70.
7. Loeb S, Bjurlin MA, Nicholson J, et al. Overdiagnosis and Overtreatment of Prostate Cancer. *European urology* 2014.
8. Bul M, van den Bergh RC, Zhu X, et al. Outcomes of initially expectantly managed patients with low or intermediate risk screen-detected localized prostate cancer. *BJU international* 2012; 110(11): 1672-7.
9. Cooperberg MR, Cowan JE, Hilton JF, et al. Outcomes of active surveillance for men with intermediate-risk prostate cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011; 29(2): 228-34.
10. Klotz L. Prostate cancer overdiagnosis and overtreatment. *Current opinion in endocrinology, diabetes, and obesity* 2013; 20(3): 204-9.
11. Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine* 2012; 157(2): 120-34.
12. Heidenreich A, Abrahamsson PA, Artibani W, et al. Early detection of prostate cancer: European Association of Urology recommendation. *European urology* 2013; 64(3): 347-54.
13. Dahabreh IJ, Chung M, Balk EM, et al. Active surveillance in men with localized prostate cancer: a systematic review. *Annals of internal medicine* 2012; 156(8): 582-90.
14. Lecornet E, Ahmed HU, Hu Y, et al. The accuracy of different biopsy strategies for the detection of clinically important prostate cancer: a computer simulation. *The Journal of urology* 2012; 188(3): 974-80.
15. Taira AV, Merrick GS, Bennett A, et al. Transperineal template-guided mapping biopsy as a staging procedure to select patients best suited for active surveillance. *American journal of clinical oncology* 2013; 36(2): 116-20.
16. Kobus T, Vos PC, Hambrock T, et al. Prostate cancer aggressiveness: in vivo assessment of MR spectroscopy and diffusion-weighted imaging at 3 T. *Radiology* 2012; 265(2): 457-67.
17. Selnaes KM, Heerschap A, Jensen LR, et al. Peripheral zone prostate cancer localization by multiparametric magnetic resonance at 3 T: unbiased cancer identification by matching to histopathology. *Investigative radiology* 2012; 47(11): 624-33.
18. Turkbey B, Pinto PA, Mani H, et al. Prostate cancer: value of multiparametric MR imaging at 3 T for detection--histopathologic correlation. *Radiology* 2010; 255(1): 89-99.
19. Ahmed HU, Kirkham A, Arya M, et al. Is it time to consider a role for MRI before prostate biopsy? *Nature reviews Clinical oncology* 2009; 6(4): 197-206.
20. Moore CM, Robertson NL, Arsanious N, et al. Image-guided prostate biopsy using magnetic resonance imaging-derived targets: a systematic review. *European urology* 2013; 63(1): 125-40.

21. Rastinehad AR, Turkbey B, Salami SS, et al. Improving Detection of Clinically Significant Prostate Cancer: MRI/TRUS Fusion-Guided Prostate Biopsy. *The Journal of urology* 2013.

22. Haffner J, Lemaitre L, Puech P, et al. Role of magnetic resonance imaging before initial biopsy: comparison of magnetic resonance imaging-targeted and systematic biopsy for significant prostate cancer detection. *BJU international* 2011; 108(8 Pt 2): E171-8.

23. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *European radiology* 2012; 22(4): 746-57.

24. Kuru TH, Roethke MC, Rieker P, et al. Histology core-specific evaluation of the European Society of Urogenital Radiology (ESUR) standardised scoring system of multiparametric magnetic resonance imaging (mpMRI) of the prostate. *BJU international* 2013; 112(8): 1080-7.

25. Portalez D, Mozer P, Cornud F, et al. Validation of the European Society of Urogenital Radiology scoring system for prostate cancer diagnosis on multiparametric magnetic resonance imaging in a cohort of repeat biopsy patients. *European urology* 2012; 62(6): 986-96.

26. Moore CM, Kasivisvanathan V, Eggener S, et al. Standards of Reporting for MRI-targeted Biopsy Studies (START) of the Prostate: Recommendations from an International Working Group. *European urology* 2013.

27. Wolters T, Roobol MJ, van Leeuwen PJ, et al. A critical analysis of the tumor volume threshold for clinically insignificant prostate cancer using a data set of a randomized screening trial. *The Journal of urology* 2011; 185(1): 121-5.

28. Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med* 2007; 26(10): 2170-83.

29. Stamey J, Holt, M.M. Bayesian interval estimation for predictive values from case-control studies. *Communications in Statistics - Simulation and Computation* 2010; 39(1): 101-10.

30. Enoe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med* 2000; 45(1-2): 61-81.

31. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics* 1980; 36(1): 167-71.

32. Hambrock T, Somford DM, Hoeks C, et al. Magnetic resonance imaging guided prostate biopsy in men with repeat negative biopsies and increased prostate specific antigen. *The Journal of urology* 2010; 183(2): 520-7.

33. 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. *The Journal of urology* 2013; 189(3): 860-6.

34. Van de Ven WJ, Barentsz J. MRI/US-guided biopsy - a viable alternative to TRUS-guidance. *Nature reviews Urology* 2013; In press.

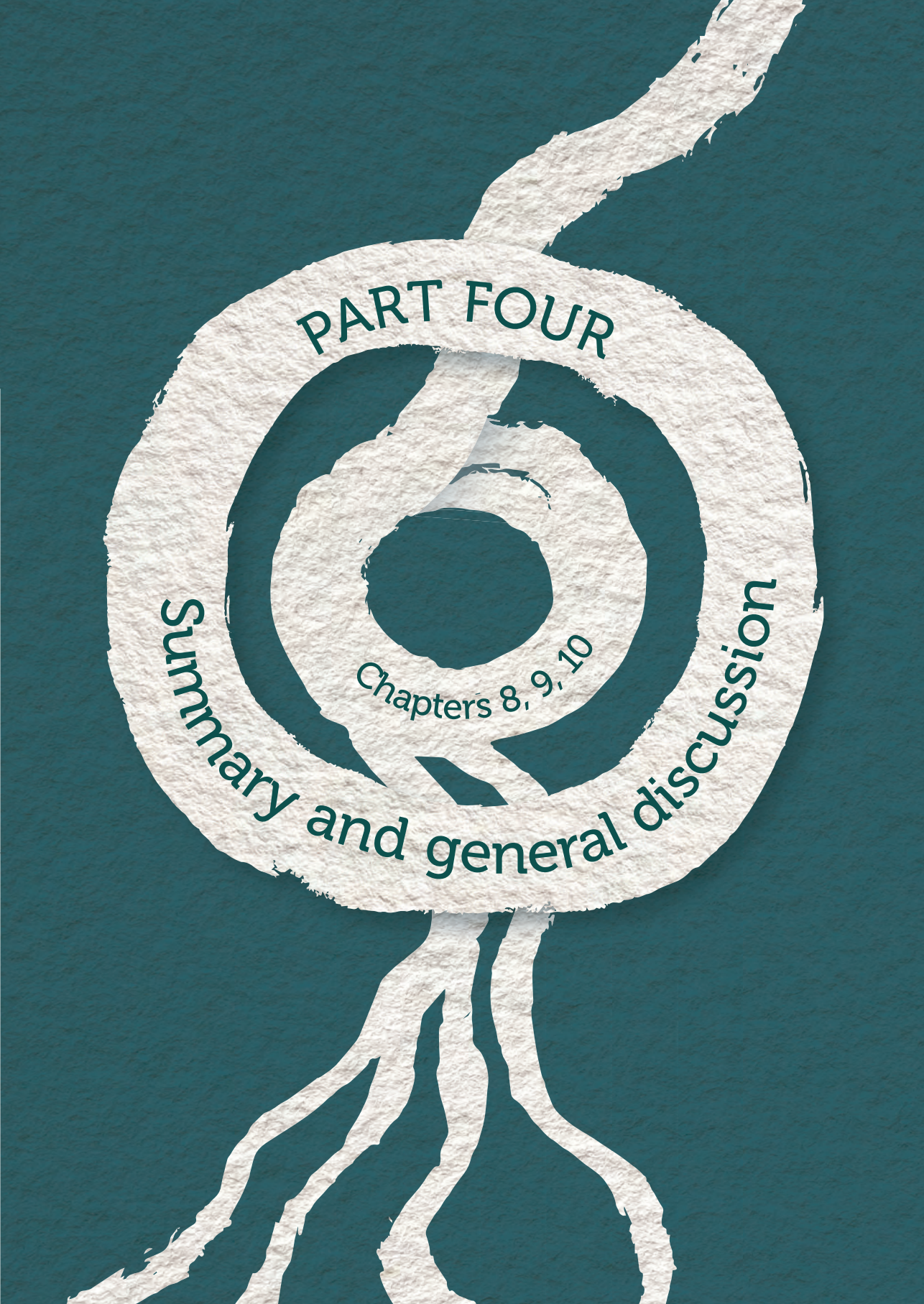
35. Siddiqui MM, Rais-Bahrami S, Truong H, et al. Magnetic resonance imaging/ultrasound-fusion biopsy significantly upgrades prostate cancer versus systematic 12-core transrectal ultrasound biopsy. *European urology* 2013; 64(5): 713-9.

36. van de Ven WJ, Hulsbergen-van de Kaa CA, Hambrock T, Barentsz JO, Huisman HJ. Simulated required accuracy of image registration tools for targeting high-grade cancer components with prostate biopsies. *European radiology* 2013; 23(5): 1401-7.

37. de Rooij M, Crienen S, Witjes JA, Barentsz JO, Rovers MM, Grutters JP. Cost-effectiveness of Magnetic Resonance (MR) Imaging and MR-guided Targeted Biopsy Versus Systematic Transrectal Ultrasound-Guided Biopsy in Diagnosing Prostate Cancer: A Modelling Study from a Health Care Perspective. *European urology* 2013.

38. Schroder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *The New England journal of medicine* 2012; 366(11): 981-90.

39. Schroder FH, Hugosson J, Carlsson S, et al. Screening for prostate cancer decreases the risk of developing metastatic disease: findings from the European Randomized Study of Screening for Prostate Cancer (ERSPC). *European urology* 2012; 62(5): 745-52.


PART FOUR
Chapters 8, 9, 10
Summary and general discussion

8

General discussion and future directions

New MRI technologies might be the answer to counterbalance the limitations of the current prostate cancer (PCa) diagnostic pathway. However, high-quality evidence regarding the superiority of these new technologies in terms of accuracy, efficacy, and (cost-)effectiveness is needed to cause a paradigm shift, and to allow their possible implementation.

This aim of this thesis was to evaluate the diagnostic multiparametric MRI (mp-MRI) pathway in men with a suspicion of PCa. More specifically, this thesis aimed to assess the diagnostic and staging accuracy of mp-MRI, and the (cost-)effectiveness of an image based pathway. Our diagnostic meta-analysis in **Chapter 2**, showed pooled sensitivity and specificity of 0.74 and 0.88 for detection of all types of PCa using mp-MRI. In our diagnostic meta-analysis in **Chapter 3**, pooled sensitivity for mp-MRI for detection of PCa improved when only studies were included that reported on clinically significant PCa only (0.74 versus 0.84), with specificity values that decreased (from 0.80 to 0.75). Furthermore, accuracy seemed to improve when the Prostate Imaging Reporting and Data System (PI-RADS) was used accurately. Our prospective study in **Chapter 7**, showed that a new image-based pathway using mp-MRI can improve the diagnosis of clinically significant PCa (increase of 18%), reduce over-diagnosis of indolent PCa with subsequent over-treatment (reduction of 89%), and might reduce the number of unnecessary prostate biopsies (reduction of 51%). Moreover, an MRI driven diagnostic pathway seemed to be cost-effective compared to the standard of care, with higher overall quality of life (in QALYs) for an acceptable price (**Chapter 5**). For local staging of PCa, mp-MRI has poor and heterogeneous sensitivity values (extracapsular extension (ECE) 0.57; seminal vesicle invasion (SVI) 0.58; overall stage T3 0.61) with high specificity values (ECE 0.91; SVI 0.96; T3 0.88), but sensitivity seemed to slightly improve when higher magnetic field strengths and multiparametric MR techniques were used (**Chapter 4**).

However, it should be acknowledged that a technology like mp-MRI is not just an isolated device, but something that is interwoven within a societal and clinical context. Involvement of all relevant stakeholders will lead to a better understanding of what a technology actually does in a societal context. Furthermore, an overview of all stakeholders will open large sources of relevant expertise and thus generates potentially important information, ie, it will help to identify a comprehensive list of perspectives and issues related to a technology. In this discussion, we will not present a formal stakeholder analysis, but we will describe the most important challenges for the different stakeholders involved following from the most important findings from this thesis. Hopefully, such an overview results in creative future directions regarding the implementation of the technique.

STAKEHOLDERS IN PROSTATE CANCER DIAGNOSTIC PATHWAY

Although the patient and his urologist are the main actors in the PCa diagnostic pathway, many other stakeholders are involved, such as his relatives, general practitioner (GP), other medical specialists, and policy makers (see also Figure 1 and Table 1). In general, the patient and his relatives have to cope with the uncertainty in the diagnostic work-up, and the discomfort after possible treatment. His GP needs to have state-of-art knowledge about the recommendations on the role of prostate specific antigen (PSA) screening and the consequences of the test results. His urologist must be able to inform the patient about the available diagnostic and treatment options, and make a shared decision. Ideally, the urologist is supported in this decision-making process through a multidisciplinary approach, in which radiologists, pathologists, and (medical and radiation) oncologists are consulted to be optimally informed about the aggressiveness, location and invasiveness of the tumour, and its possible treatment options. The clinical framework in which the GP, the urologist, and other medical specialists operate is provided by guideline developers and health care insurers.

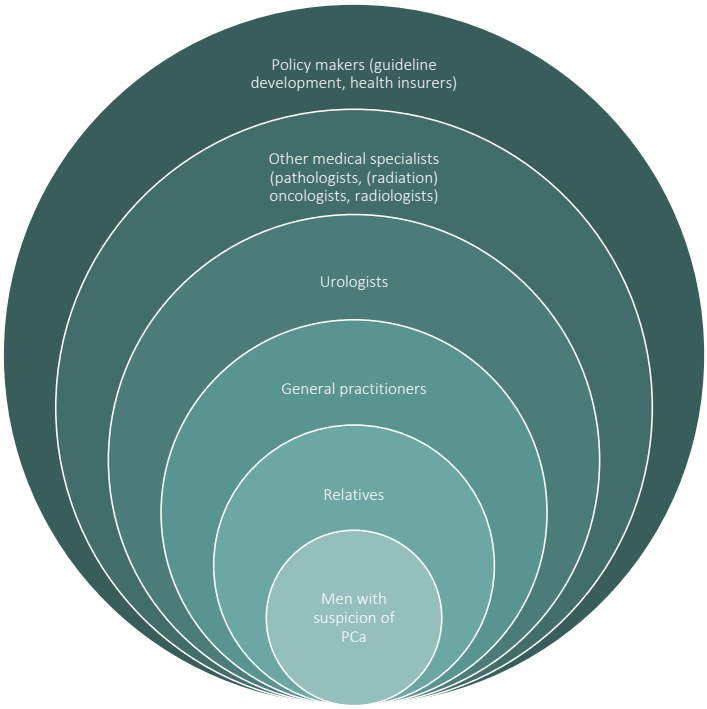


Figure 1 Different stakeholders in the diagnostic pathway of prostate cancer.

CHALLENGES FOR STAKEHOLDERS AND EVIDENCE FROM THIS THESIS

Patient and relatives

The diagnostic pathway that men with a suspicion of PCa undergo is known to be a stressful period. It is particularly important for patients and their relatives that the time of uncertainty will be as short as possible, that the pathway will harbour an accurate diagnosis, and that the diagnostic procedures will be as minimally invasive as possible. In the current diagnostic pathway, all men with a suspicion of PCa undergo systematic 10-12 core transrectal ultrasound guided biopsy (TRUS-GB); an invasive test with a low accuracy. This thesis shows that an MRI based diagnostic pathway in which an elevated PSA is followed by mp-MRI, and if necessary, MR targeted prostate biopsy, can prevent unnecessary biopsies and improve diagnostic accuracy (see *Chapter 7*). Other studies show that MR guided in-bore biopsy (MRGB) of the prostate seems to have fewer side effects and lower pain severity scores than TRUS-GB, and was preferred by most patients.[1]

General practitioner

Since the introduction of a PSA driven diagnostic pathway, more men with early PCa are identified, which facilitated curative treatment at an early stage in the disease history. However, it has caused multiple problems that have been addressed in a series of taskforces and publications. The U.S. Preventive Services Task Force (USPSTF) recently recommended against widespread use of PSA as screening test due to the potential disadvantages of the 'next steps' after an elevated PSA (TRUS-GB), which may cause harm and might not increase quality and length of life.[2] Furthermore, the Dutch PCa guideline for GPs recommend against active PSA testing, and testing should be done only after careful consideration of the (dis-)advantages of the test results.[3] Despite this caution, PSA testing, although less commonly than before, still occurs, mainly because men ask for active PSA screening.[4,5] The disadvantages of the random TRUS-GB pathway are well known, and it continues to find low-risk disease. The mp-MRI as triage test after an elevated PSA and/or abnormal DRE can potentially improve the PSA dilemma and reduce over-diagnosis and over-treatment because it enables better stratification between indolent and aggressive PCa (see *Chapter 7*).

Ideally, a minimally invasive screening/triage test, that is less time consuming, and less expensive than the combination of PSA with prostate MRI should be developed. Several new tests are being under investigation, to possibly replace or improve the current PSA screening, like molecular biomarkers, genetic markers and modified PSA testing. However,

it is difficult to ascertain which biomarker is the most reliable among the available tests. A personalized screening based on a combination of PSA testing, individual genetic profile, mp-MRI, and other biomarkers, can potentially reduce the number of unnecessary biopsies and increase the identification of clinically significant disease and helps to identify the most optimal screening parameters.[6]

Urologists

The optimal diagnostic pathway for PCa should be able to identify men with clinically significant PCa, while avoiding detection of indolent disease. Because the current standard of care is associated with inaccurate diagnosis, and the identification of new molecular biomarkers that identify significant PCa is not robust enough for use in individual men, urologists increasingly incorporate new imaging techniques into the diagnostic pathway. Among the different new imaging techniques (eg, enhanced TRUS and advanced MRI techniques), mp-MRI is currently the most promising and best studied new imaging technique.

Pooled sensitivity and specificity values of mp-MRI that are presented in the diagnostic meta-analysis in *Chapter 2* of this thesis (sensitivity 0.74 (95% CI 0.67-0.92); specificity 0.88 (95% CI 0.82-0.92)) are based on the detection of all PCa types; ie, also for the detection of clinically insignificant tumours. Although these results are promising for improved detection, urologists are more interested in the ability of mp-MRI to distinguish indolent from aggressive disease. That is why in more recent meta-analyses the role of mp-MRI in the detection of clinically significant versus indolent PCa was assessed,[7,8] similar to our meta-analysis in *Chapter 3*. These studies showed that the presence of a tumour suspicious area on mp-MRI in men with low-risk PCa on TRUS-GB results in a high likelihood that clinically significant PCa will be present. On the other hand, if a lesion is not seen on mp-MRI, the attribution of low-risk PCa is much more likely to be correct.[9] A number of recent systematic reviews and meta-analyses have consistently shown that targeted biopsy strategies using predefined lesions identified by mp-MRI have better accuracy and efficiency compared with TRUS-GB,[10,11] similar to our findings in *Chapter 7*.

Although the exact recommendations regarding the use of mp-MRI and MR targeted biopsy in PCa are yet to be defined, there are four important patient categories that could clearly benefit from the use of an mp-MRI driven diagnostic pathway;

(1) Biopsy-naive men with a suspicion of PCa based on an elevated PSA and/or positive digital rectal examination (DRE). An MRI based pathway can be helpful to

triage men with a suspicion of PCa. In men with nonsuspicious MRI findings, the likelihood of clinically significant disease appears to be sufficiently low to defer invasive treatment, and consider PSA follow-up without further confirmatory biopsy, as shown in the prospective study in *Chapter 7*.[12] However, because of the lack of longer follow-up, we do not exactly know how many men might safely avoid a biopsy as a result of negative imaging. Studies are currently being performed to address this issue.[13]

(2) In men with previous negative prostate biopsy, and persistent clinical suspicion, mp-MRI may point to a lesion that may have been missed on previous TRUS-GB or provide an explanation for a rising PSA in the absence of cancer (e.g. prostatitis or BPH).[14]

(3) Men with biopsy-proven PCa. (mp-)MRI has been reported to be effective for detection of extra-prostatic disease (ECE, SVI), and lymph node involvement. Although the sensitivity values of (mp-)MRI for local staging are still poor to moderate (as shown in *Chapter 4* of this thesis), mp-MRI may help in treatment planning, including nerve-sparing surgery, radiation therapy, and focal therapy. Nerve sparing prostatectomy may preserve erectile function in a considerable proportion of patients.[15] The ability of mp-MRI to localize and map prostatic tumour as part of the pre-operative evaluation can aid greatly in the planning of nerve sparing techniques.[16,17] Superior anatomical resolution of mp-MRI permits the detection of clinically significant PCa lesions, which has also enabled image guided focal therapy of PCa with minimally invasive techniques, such as cryosurgical ablation, high intensity focused ultrasound (HIFU), laser evaporation, and irreversible electroporation.[18] Mp-MRI can also be used for the planning of radiation therapy.[19]

(4) Men with a suspicion of recurrence after previous definitive treatment of PCa. In men with a suspicion of recurrence, mp-MRI can help to identify the site of recurrence.[20]

Radiologist

Radiologists strive for the same goal as urologists: they also aim to identify PCa that needs treatment, while trying to avoid over-diagnosis and over-treatment of indolent disease. While urologists focus specifically on the clinical aspects, radiologists are engaged in the

more technical aspects of this clinical problem. Their main concerns are optimization of the existing MRI protocols for the detection and local staging of PCa and development of a standardized and structured form of reporting to communicate their findings. These 'technical' aspects are essential to implement before further incorporation of prostate MRI into standard disease management.

The question remains what the optimal scanning protocol is. The mp-MRI should be accurate, reproducible and fast. Minimal technical requirements and optimal scanning protocols for detection and local staging of PCa are defined by the European Society for Urogenital Radiology (ESUR)[21] based on current literature and expert opinion. For the detection of PCa the use of at least two additional techniques (diffusion weighted imaging (DWI), dynamic contrast-enhanced MRI (DCE-MRI), and/or magnetic resonance spectroscopic imaging (MRSI) is recommended. Higher magnetic field strengths (preferably 3.0 instead of 1.0 or 1.5 Tesla) lead to higher accuracy.[22] Although there is consensus about the added value of using mp-MRI techniques over standard anatomic MRI only, the best combination still needs to be elucidated. The 'full package', using T2WI with additional DWI, DCE-MRI, and MRSI should be the best possible combination. The question can be raised whether all these techniques are really necessary for the detection of clinically significant PCa. The answer is probably "no"; the optimal combination is highly dependent on the clinical context. For screening purposes, some advocate the use of anatomic T2-weighted imaging (T2WI) in combination with DWI only, which can be performed quickly, at low cost.[23] In groups with higher pre-test probabilities on PCa, for example in men with persistent elevated PSA, adding DCE-MRI or MRSI sequences with or without an endorectal coil, could possibly improve accuracy.[24] The aim of this thesis was not to find the best MRI protocol for detection of PCa. However, for local staging of PCa, we tried to identify technical (MRI) characteristics that resulted in the highest accuracy for local staging of PCa (*Chapter 4*). Especially higher magnetic field strengths (3.0 Tesla) and mp-MRI techniques seemed to slightly improve sensitivity values. Although individual studies showed improved local staging accuracy with the use of an endorectal coil,[25] we could not demonstrate this proposed added value in our local staging meta-analysis (*Chapter 4*).

Equally important to a standardized scanning protocol are the reliability and reproducibility of the reader's interpretation of the mp-MRI. To harmonize reporting prostate MRIs, the ESUR introduced the Prostate Imaging Reporting and Data System (PI-RADS) in 2012.[21] This system aims to standardize imaging acquisition and reporting, in analogy with the Breast Imaging Reporting and Data System (BI-RADS). PI-RADS uses a 5-point scale, ranging

from highly unlikely (score of 1) to highly likely (score of 5) that clinically significant PCa is present. The modified version of PI-RADS is recently published in 2015 in collaboration with the American College of Radiology (ACR) and the AdMeTech Foundation.[26] In our diagnostic meta-analysis in *Chapter 3* fourteen studies (1785 patients) could be analysed, with pooled sensitivity of 0.78 (95% CI 0.70-0.84) and specificity of 0.79 (95% CI 0.68-0.86) for PCa detection. In studies with the correct use of PI-RADS (version 1, 2012), higher sensitivity and specificity were found compared with studies with a less strict or adjusted use of the PI-RADS criteria.

Policy makers

Policy makers (such as health insurance companies and guideline developers) are striving for an affordable and high quality evidence-based health care for men with (a suspicion of) PCa, while minimizing the risk of over-diagnosis and over-treatment. As mentioned above, the current PSA based screening has produced modest reduction in PCa mortality, but this has come at the cost of a substantial increase in the detection of men with low-risk PCa.[27] Unless the indolent nature and the long-life expectancy of the affected men, they often choose for an active treatment of these cancers. Over-diagnosis and over-treatment of this indolent tumours leads to unnecessary side effects and has cost a vast sum of precious health care money.[6] Although it is widely accepted that mp-MRI can improve the diagnostic pathway of PCa, the role of mp-MRI in PCa guidelines is still limited. European and Dutch PCa guidelines recommend mp-MRI in men with a remaining suspicion of clinical significant PCa after previous negative TRUS-GB.[28] A recent meta-analysis in this group of men showed that targeted MR guided biopsy could increase the detection of significant PCa while the number of men with insignificant cancer could be reduced.[7] In biopsy-naive men with an abnormal PSA or DRE, systematic TRUS-GB is still recommended as the next diagnostic procedure, unless there is increasing evidence that diagnosis of insignificant PCa and thus over-treatment of PCa could be drastically reduced with an mp-MRI driven diagnostic pathway.

The limited role of mp-MRI in clinical guidelines is one of the reasons that health insurance companies still do not reimburse mp-MRI for biopsy-naive men. Another possible reason is that the cost-effectiveness of an imaging pathway needs to be examined, and that the technique is currently only available in expert centres. In *Chapter 6* of this thesis, we showed that an MRI pathway seems to be cost-effective compared to the standard TRUS-GB pathway in men with a suspicion of PCa. A handful of other studies also assessed the cost-effectiveness of an MRI based diagnostic pathway, which do not provide a definitive

answer as to whether mp-MRI is cost-effective.[29] This difference can be explained by important differences in the diagnostic procedures, patient groups (for example biopsy-naive men versus previous negative TRUS-GB), and the assumptions incorporated in the different modelling structures. Another possible explanation of the limited role of mp-MRI in clinical guidelines is the lack of evidence of superiority compared with the standard of care in terms of patient outcomes, which has to be addressed in future research.

FUTURE DIRECTIONS

In the recent years, the role of mp-MRI has evolved from add-on test after negative TRUS-GB to rather a potential triage/screening test in men with an elevated PSA, followed by targeted biopsy.[9] However, there are several obstacles to realize the potential benefits of the proposed image-based pathway:[30]

(1) To disseminate an mp-MRI pathway in clinical practice, it should be further standardized regarding the technical equipment, examination protocols, image acquisition, processing, and post-processing.

(2) Training, quality criteria and certification. The availability of an adequate MRI device is for most centres not an issue, but radiologists must be able to interpret prostate MRI, and communicate with their urologists in a standardized lexicon (PI-RADS). For training and quality purposes, communication between low-volume centres and high-volume centres should be promoted, with training programs and double-reading. Just like in mammography, quality control of imaging and reporting will be required to ensure centres do not over-call or under-call suspicious lesions. Together with policy makers, the radiological and urological community need to draft quality criteria in order to provide a high quality clinical diagnosis.

(3) The medical and societal community must fully acknowledge the distinct nature of low-risk PCa as an entity that does not require active treatment.

(4) Greater pressure and demand by patients and physicians, as well updated national PCa guidelines that endorse mp-MRI and MR targeted biopsy, could influence payers to update their policies and centralize care in expert centres.

CONCLUSIONS

In this thesis, we showed that mp-MRI improves the diagnostic pathway in a cost-effective manner in biopsy-naive men by reducing the diagnosis of low-risk PCa, and improving the detection of clinically significant PCa. Therefore, it seems justified that the Dutch Prostate Cancer Foundation endorse the introduction of an mp-MRI based pathway for men with a suspicion on PCa. So, more important is the question: how to proceed?

The overall goal of all stakeholders is to perform biopsies and to treat only those men for whom treatment will result in a reduced mortality and higher quality of life, while lowering the burden of an invasive diagnostic pathway from those who will not benefit. Ideally, this diagnostic pathway takes place in a centre with state-of-art diagnostic options and high quality of care, independent from which centre you visit. A prerequisite to succeed is multidisciplinary centralized care in expert centres, in which urologists, radiologists, pathologists, and (radiation and medical) oncologists work together to provide men the best available clinical diagnosis and treatment advice.

Challenges for different stakeholders	Evidence from this thesis	Gaps in knowledge / future directions
Patients and relatives		
Burden of prostate biopsies	An mp-MRI pathway can reduce the number of unnecessary biopsies (by at least 50%) and allows for targeted biopsy procedures	Need for evaluation of best targeted biopsy technique (in-bore, fusion, cognitive) for the identified tumour suspicion lesions
Uncertainty around diagnosis and local staging	An mp-MRI pathway can improve the detection of clinically significant PCa Mp-MRI and can be helpful in identification of local tumour growth outside the prostate, but sensitivity is poor	Need for identification of characteristics of missed tumours, since not all clinically significant tumours are identified with mp-MRI Need for improved sensitivity of mp-MRI for local staging of PCa Need for studies on QoL of men in mp-MRI diagnostic pathway
Over-diagnosis and over-treatment of PCa	An mp-MRI pathway can reduce over-diagnosis and over-treatment because indolent tumours are mostly not seen on mp-MRI	Need for better characterization of missed clinically significant tumours (under-diagnosis)
General practitioners		
Role of PSA test and alternatives, and 'next' steps after an elevated PSA	Mp-MRI shows potential to be used as triage test after an elevated PSA and seems to be cost-effective In this thesis, the role of mp-MRI as screening test -with or without PSA- is not studied	Need for evaluation of mp-MRI as screening test after elevated PSA Need for more accurate low cost first line tests instead of low specific PSA test (e.g. new biomarker tests) Need for knowledge by the GP what to do with patients with an elevated PSA and a negative mp-MRI (education by the urologists)
Urologists		
Accuracy of biopsy technique	An mp-MRI pathway reduces over-diagnosis of indolent disease and improves detection of clinically significant prostate cancer	Need for identification of optimal biopsy protocol; is it possible to use targeted biopsy only and if yes, what type of targeted biopsy technique?
Post-biopsy complications	An mp-MRI pathway can reduce the number of unnecessary biopsies and can thus reduce the overall number of post-biopsy complications	Need for assessment of risk on complications of standard 10-12 core TRUS-GB versus 2-4 core targeted biopsy

Accuracy of detection of clinical significant PCa	Mp-MRI improves the detection of clinical significant PCa compared to standard of care	Need for studies with longer follow-up to evaluate 'negative' mp-MRI cases and assess 'real' mp-MRI accuracy
Accuracy of local staging of PCa	Mp-MRI shows high specificity but poor sensitivity for local staging of PCa Higher field strengths (i.e. 3.0T) and mp-MRI seems to improve sensitivity	Need for evaluation of the role of local staging mp-MRI; high sensitivity or high specificity more important for pre-treatment planning and NVB?
Radiologists		
Protocol for diagnostic mp-MRI	Mp-MRI shows high but variable sensitivity and NPV and high specificity for the detection of PCa, which implies a potential role before biopsy	Need for identification of most optimal combination of different mp-MRI techniques. For screening, there is need for fast, low cost, bi-parametric MRI
Protocol and reading of local staging mp-MRI	For local staging, mp-MRI shows high specificity, but poor sensitivity, which can be improved by higher magnetic field strength and the use of mp-MRI techniques	Need for further evaluation of the role of better techniques in the MRI protocol for local staging
International standardization of reporting prostate MRI	The proper use of PI-RADS as standardized reporting system for mp-MRI seems to improve accuracy for PCa detection	Need for education programs, set up of minimal quality criteria to only allow specialized centres to perform mp-MRI, set up of an independent quality control program (quality assessment), and certification before standard utilization of mp-MRI for detection and local staging of PCa
Policy makers (guidelines / health insurance)		
Cost-effectiveness of an MRI driven diagnostic pathway	An mp-MRI pathway seems to be cost-effective compared to the standard of care using TRUS-GB	Need for evaluation of possibility to implement mp-MRI for detection of PCa on a larger/national scale
Role of mp-MRI in guideline for detection of PCa	In biopsy naïve men an mp-MRI pathway can reduce unnecessary biopsies and diagnosis of indolent disease, and improve detection of clinically significant PCa For men with previous TRUS-GB and persistent clinical suspicion, mp-MRI shows high sensitivity and specificity for PCa detection	Need for assessment of optimal diagnostic pathway in multidisciplinary discussions using the most recent evidence
Role of mp-MRI in guideline for local staging of PCa	For local staging mp-MRI shows high specificity, but poor sensitivity, which can be improved by higher magnetic field strength and the use of mp-MRI techniques	Need for evaluation of other imaging techniques and updated guidelines
Quality assurance of mp-MRI	The techniques and standardized reporting can be learned by radiologists and urologist but has a learning curve	International certified training, set up of quality assessment criteria, and quality control

REFERENCES

1. Egbers N, Schwenke C, Maxeiner A, Teichgraber U, Franiel T. MRI-guided core needle biopsy of the prostate: acceptance and side effects. *Diagnostic and interventional radiology* 2015; 21(3): 215-21.
2. Moyer VA, Force USPST. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine* 2012; 157(2): 120-34.
3. Blanker MH, Klomp MA, van den Donk M, van der Heide WK, Opstelten W, Burgers JS. Samenvatting van de NHG-standaard 'Mictieklachten bij mannen'. *Ned Tijdschr Geneeskd* 2013; 157(18): A6178.
4. Jemal A, Fedewa SA, Ma J, et al. Prostate Cancer Incidence and PSA Testing Patterns in Relation to USPSTF Screening Recommendations. *Jama* 2015; 314(19): 2054-61.
5. Hamoen EH, Reukers DF, Numans ME, Barentsz JO, Witjes JA, Rovers MM. Discrepancies between guidelines and clinical practice regarding prostate-specific antigen testing. *Family practice* 2013; 30(6): 648-54.
6. Gandaglia G, Briganti A, Fossati N, et al. The Problem Is Not What to Do with Indolent and Harmless Prostate Cancer-The Problem Is How to Avoid Finding These Cancers. *Eur Urol* 2016; 70(4): 547-8.
7. Schoots IG, Roobol MJ, Nieboer D, Bangma CH, Steyerberg EW, Hunink MG. Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. *Eur Urol* 2015; 68(3): 438-50.
8. Futterer 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(6): 1045-53.
9. Valerio M, Willis S, van der Meulen J, Emberton M, Ahmed HU. Methodological considerations in assessing the utility of imaging in early prostate cancer. *Current opinion in urology* 2015; 25(6): 536-42.
10. Valerio M, Donaldson I, Emberton M, et al. Detection of Clinically Significant Prostate Cancer Using Magnetic Resonance Imaging-Ultrasound Fusion Targeted Biopsy: A Systematic Review. *Eur Urol* 2015; 68(1): 8-19.
11. Moore CM, Robertson NL, Arsanious N, et al. Image-guided prostate biopsy using magnetic resonance imaging-derived targets: a systematic review. *Eur Urol* 2013; 63(1): 125-40.
12. Abd-Alazeez M, Ahmed HU, Arya M, et al. Can multiparametric magnetic resonance imaging predict upgrading of transrectal ultrasound biopsy results at more definitive histology? *Urologic oncology* 2014; 32(6): 741-7.
13. Parker C, Kaplan RS, Gabe R, et al. Multiparametric magnetic resonance imaging in the diagnosis and characterisation of prostate cancer: The PROMIS study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012; 30(5_suppl): 62.
14. Durmus T, Reichelt U, Huppertz A, Hamm B, Beyersdorff D, Franiel T. MRI-guided biopsy of the prostate: correlation between the cancer detection rate and the number of previous negative TRUS biopsies. *Diagnostic and interventional radiology* 2013; 19(5): 411-7.
15. Finley DS, Margolis D, Raman SS, et al. Fine-tuning robot-assisted radical prostatectomy planning with MRI. *Urologic oncology* 2013; 31(6): 766-75.
16. Park BH, Jeon HG, Jeong BC, et al. Influence of magnetic resonance imaging in the decision to preserve or resect neurovascular bundles at robotic assisted laparoscopic radical prostatectomy. *The Journal of urology* 2014; 192(1): 82-8.
17. Roethke MC, Lichy MP, Kniess M, et al. Accuracy of preoperative endorectal MRI in predicting extracapsular extension and influence on neurovascular bundle sparing in radical prostatectomy. *World journal of urology* 2013; 31(5): 1111-6.
18. Valerio M, Cerantola Y, Eggener SE, et al. New and Established Technology in Focal Ablation of the Prostate: A Systematic Review. *Eur Urol* 2017; 71(1): 17-34.

19. McPartlin AJ, Li XA, Kershaw LE, et al. MRI-guided prostate adaptive radiotherapy - A systematic review. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 2016; 119(3): 371-80.
20. Oppenheimer DC, Weinberg EP, Hollenberg GM, Meyers SP. Multiparametric Magnetic Resonance Imaging of Recurrent Prostate Cancer. *Journal of clinical imaging science* 2016; 6: 18.
21. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *European radiology* 2012; 22(4): 746-57.
22. Lagemaat MW, Scheenen TW. Role of high-field MR in studies of localized prostate cancer. *NMR in biomedicine* 2014; 27(1): 67-79.
23. De Visschere P, Lumen N, Ost P, Decaestecker K, Pattyn E, Villeirs G. Dynamic contrast-enhanced imaging has limited added value over T2-weighted imaging and diffusion-weighted imaging when using PI-RADSv2 for diagnosis of clinically significant prostate cancer in patients with elevated PSA. *Clinical radiology* 2016.
24. Vos EK, Litjens GJ, Kobus T, et al. Assessment of prostate cancer aggressiveness using dynamic contrast-enhanced magnetic resonance imaging at 3 T. *Eur Urol* 2013; 64(3): 448-55.
25. Futterer JJ, Engelbrecht MR, Jager GJ, et al. Prostate cancer: comparison of local staging accuracy of pelvic phased-array coil alone versus integrated endorectal-pelvic phased-array coils. Local staging accuracy of prostate cancer using endorectal coil MR imaging. *European radiology* 2007; 17(4): 1055-65.
26. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* 2016; 69(1): 16-40.
27. Hayes JH, Barry MJ. Screening for prostate cancer with the prostate-specific antigen test: a review of current evidence. *Jama* 2014; 311(11): 1143-9.
28. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* 2016.
29. Willis SR, van der Meulen J, Valerio M, Miners A, Ahmed HU, Emberton M. A review of economic evaluations of diagnostic strategies using imaging in men at risk of prostate cancer. *Current opinion in urology* 2015; 25(6): 483-9.
30. Rosenkrantz AB, Taneja SS. Prostate MRI can reduce overdiagnosis and overtreatment of prostate cancer. *Academic radiology* 2015; 22(8): 1000-6.

9

Summary

Prostate cancer (PCa) is the most common malignancy among men in developed countries, and has a large societal burden. Approximately 50-70 % of the detected PCa is a low-risk type that will never cause symptoms during a lifetime. However, in the other 30-50% of the men, an intermediate- or high-risk type is diagnosed, which needs to be treated. The challenge is to identify the aggressive tumours early and to treat these accordingly, while leaving alone low-risk PCa.

Systematic transrectal ultrasound guided biopsy (TRUS-GB) is the current standard of care after an elevated serum prostate specific antigen (PSA). It is an invasive procedure, and it is known that important PCa is being missed or misclassified with this method, whereas low-risk PCa which does not require treatment, may be found. This leads to over-diagnosis and over-treatment of low-risk tumours, and under-treatment of intermediate- or high-risk PCa. Multiparametric magnetic resonance imaging (mp-MRI), preferably used as an advanced imaging technique before biopsy, shows promising results in improving PCa diagnosis. However, mp-MRI before biopsy is not recommended in the national PCa guidelines, and many health insurance companies withhold reimbursement for mp-MRI as (first) diagnostic step after an elevated PSA. There are several factors that may be an explanation for this: the evidence regarding the effectiveness of mp-MRI on quality of life and its related cost-effectiveness is lacking, and there is a lack of evidence of superiority of a diagnostic MRI pathway over a TRUS-GB pathway in terms of patient outcome. Furthermore, mp-MRI might not be available in all hospitals.

This thesis aimed to evaluate the diagnostic MRI pathway in men with a suspicion of PCa. Therefore, we studied the diagnostic accuracy of mp-MRI for the detection and local staging of PCa. Moreover, we assessed the clinical- and cost-effectiveness of the MRI pathway in patients with a suspicion of PCa.

In Chapter 2-4, we provide an overview of the current evidence regarding the accuracy of prostate MRI. The accuracy of mp-MRI for the detection of all types of PCa is studied in **Chapter 2.** We performed a systematic review and meta-analysis to determine the diagnostic accuracy of mp-MRI, including studies that used an MRI protocol of T2-weighted imaging (T2WI), combined with diffusion weighted imaging (DWI), and dynamic contrast enhanced MRI (DCE-MRI), based on the recommendation of the European Society of Urogenital Radiology (ESUR). Seven studies that met the inclusion criteria (526 patients) could be analysed. Pooled data showed high overall specificity of 0.88 (95% CI 0.82–0.92) and sensitivity of 0.74 (95% CI 0.66–0.81) for PCa detection, with negative predictive values

(NPVs) ranging from 0.65 to 0.94. The overall methodologic quality of the included studies was fair, but large heterogeneity was reported. Nevertheless, subgroup analyses did not show considerable differences between various subgroups.

In **Chapter 3,** we specifically studied the accuracy of mp-MRI for the detection of clinically significant PCa and the role of the Prostate Imaging Reporting and Data System (PI-RADS), which is a standardized reporting system for prostate MRI. In this diagnostic meta-analysis, fourteen studies (1785 patients) could be analysed. The pooled data showed sensitivity of 0.78 (95% CI 0.70–0.84) and specificity of 0.79 (95% CI 0.68–0.86) for clinically significant PCa detection, with NPVs ranging from 0.58 to 0.95. Sensitivity analysis showed pooled sensitivity of 0.82 (95% CI 0.72–0.89) and specificity of 0.82 (95% CI 0.67–0.92) in studies with a correct use of PI-RADS, that is, a clear description in the methodology and no adjustment of the criteria. Studies with a less strict or adjusted use of the PI-RADS criteria, or unclear description of the methodology, showed pooled sensitivity of 0.73 (95% CI 0.62–0.82) and specificity of 0.75 (95% CI 0.61–0.84). We performed another sensitivity analysis, showing slightly higher sensitivity of 0.84 (95% CI 0.76–0.89) and lower specificity of 0.75 (95% CI 0.66–0.83) in studies with detection of significant PCa as the primary outcome compared with sensitivity of 0.74 (95% CI 0.67–0.81) and specificity of 0.80 (95% CI 0.70–0.88) in studies with the detection of all PCa as the outcome measure.

In **Chapter 4,** we systematically reviewed the evidence for the use of MRI for local staging of PCa. We specifically aimed to assess the diagnostic accuracy of MRI for local staging of PCa and to analyse the influence of different imaging protocols for men with biopsy proven PCa. Radical prostatectomy specimens were used as the reference standard. We included studies that used MRI for the detection of extracapsular extension (ECE; T3a), seminal vesicle invasion (SVI; T3b), or overall stage T3 PCa. Seventy-five studies (9796 patients) could be analysed. Pooled sensitivity and specificity for ECE/T3a (45 studies, 5681 patients), SVI/T3b (34 studies, 5677 patients), and studies reporting on overall stage T3 disease only (38 studies, 4001 patients) were 0.57 (95% CI 0.49–0.64) and 0.91 (95% CI 0.88–0.93), 0.58 (95% CI 0.47–0.68) and 0.96 (95% CI 0.95–0.97), and 0.61 (95% CI 0.54–0.67) and 0.88 (95% CI 0.85–0.91), respectively. Several of the explored patient, study, and imaging characteristics, clearly influenced staging accuracy. This affected sensitivity most, whereas the specificity remained relatively stable. Higher field strengths (3.0 Tesla instead of 1.5 or 1.0 Tesla) and the use of additional functional imaging techniques (mp-MRI) seemed to improve accuracy of local staging of PCa. Use of an endorectal coil showed no additional benefit for ECE, but slightly improved the sensitivity of SVI.

The decision regarding which diagnostic strategy is preferred should not be based on diagnostic accuracy alone. In Chapter 5 and 6, we studied the costs and cost-effectiveness of the current diagnostic pathway and the new MRI pathway. In **Chapter 5,** the diagnostic meta-analysis of Chapter 2 has been used as input for a decision analytic model to assess the cost-effectiveness of the MRI strategy (mp-MRI followed by MR guided biopsy (MRGB)) versus the standard TRUS-GB strategy in diagnosing PCa. In this model, costs related to the performance and the therapeutic consequences of the tests were taken into consideration. In addition, other (in)direct consequences, such as quality of life and survival, were included in the decision analytic model. Despite the uncertainty around the presented cost-effectiveness estimates, the results of our model suggest that the MRI strategy is cost-effective in diagnosing PCa compared with the TRUS-GB strategy, assuming a sensitivity of MRGB of 20% or higher. While the MRI strategy is initially more expensive, these extra costs are compensated by reducing treatment costs because of less false positives and a better estimation of the aggressiveness of the tumour. The improvement in quality adjusted life years (QALYs) is achieved by preventing insignificant tumours being unnecessarily treated radically (with a reduced quality of life without improved survival), and reducing clinically significant PCa to be detected late (with reduced survival).

In **Chapter 6,** we provide insight in the health care costs of PCa treatment related urinary incontinence (UI) costs using a health insurance database. Reliable assessment of UI is important, because this condition is often irreversible when it still present after several years. This will have implications for the quality of life of men after treatment, but will also have major economic consequences because most men have good survival perspectives. We were able to assess UI rates and costs from longitudinal health insurance data of the Achmea Health Database (AHD), which contains representative information of approximately 17% of the Dutch population. This offered us the unique possibility to extract management strategies and UI rates with corresponding reimbursement costs. Among the different treatment options, UI rates were highest after prostatectomy. We found higher rates for laparoscopic prostatectomy compared with an open procedure, especially in the first year after treatment, which could possibly be explained by difference in experience with these techniques. The high rates and costs of UI show an estimation of the magnitude of the UI burden. This burden is likely to increase in the future, unless new strategies are adopted to improve over-diagnosis and over-treatment of PCa.

To further improve the evidence on the role of mp-MRI in the diagnostic pathway, we performed a prospective study, which is described in **Chapter 7**. We studied the clinical outcomes: the detection of low-, intermediate-, and high-risk PCa with the current standard of care and the new MRI pathway. We included 223 consecutive, biopsy-naive men, who all underwent mp-MRI and TRUS-GB. Men with equivocal or suspicious lesions on mp-MRI also underwent MRGB. The results of this diagnostic study support apparent patient benefits. Most importantly, when using the MRI pathway instead of TRUS-GB, the number of men diagnosed with low-risk PCa will be reduced by 89.4%, and reduces the need for biopsy with 51%. At the same time, the number of men diagnosed with intermediate/high-risk PCa will be increased by 17.7%.

Although the patient and his urologist are the main actors in the PCa diagnostic pathway, many other stakeholders are involved: his relatives, general practitioner, other medical specialists, and policy makers. In the **general discussion** the main findings of this thesis are discussed, and an overview is given of the most important challenges for these different stakeholders. Furthermore, we give recommendations and future directions regarding the possible implementation of the new MRI pathway.

10

Nederlandse samenvatting (Summary in Dutch)

Prostaatcarcinoom (PCa) is de meest voorkomende vorm van kanker in de Westerse wereld en heeft een grote maatschappelijke impact. Ongeveer 50-70% van de gevonden prostaattumoren zullen nooit symptomen geven, omdat dit zogenaamde laag-risico tumoren betreft. De overige 30-50% daarentegen betreffen een type tumor met een intermediair- tot hoog-risico type tumor, waarvoor het nodig is om een behandeling uit te voeren. De uitdaging is om juist deze meer agressieve tumoren zo vroeg mogelijk op te sporen en te behandelen, en de (laag-risico) tumoren die niet tot symptomen zullen leiden met rust te laten.

Volgens de huidige richtlijnen voert de uroloog systematische transrectale echogeleide prostaatbiopten (TRUS-biopten) uit bij mannen waarbij in het bloed een verhoogde prostaat specifiek antigen (PSA) waarde wordt gevonden. Een TRUS-biopt betreft een invasieve techniek met een relatief lage sensitiviteit, wat betekent dat de diagnose PCa gemist wordt, verkeerd wordt ingeschat, of dat er PCa wordt gevonden die geen behandeling behoeft. Dit kan leiden tot overdiagnose en overbehandeling van laag-risico tumoren en onderbehandeling van intermediair- of hoog-risico tumoren. Multiparametrische MRI (mp-MRI) lijkt een veelbelovend alternatief, welke de huidige PCa diagnostiek kan verbeteren, met name wanneer de techniek voorafgaand aan het nemen van prostaatbiopten wordt gebruikt. Mp-MRI voorafgaand aan prostaatbiopten wordt echter (nog) niet aanbevolen in de nationale en internationale richtlijnen voor PCa. Dit duidt erop dat de effectiviteit van mp-MRI, bijvoorbeeld op de kwaliteit van leven, alsmede de kosteneffectiviteit nog niet voldoende bewezen zijn. Daarnaast is er mogelijk onvoldoende bewijs voor de superioriteit van een diagnostisch MRI traject ten opzichte van de huidige standaardzorg waarbij TRUS-biopten worden gebruikt. Tenslotte kan het meespelen dat mp-MRI niet in alle ziekenhuizen beschikbaar is.

Het doel van het onderzoek beschreven in dit proefschrift was om het diagnostische MRI traject bij mannen met een verdenking op PCa (d.w.z. bij een verhoogd PSA en/of afwijkend rectaal toucher) te evalueren. Hiervoor hebben we de diagnostische accuratesse van mp-MRI met betrekking tot de detectie en de lokale stadiëring van PCa bestudeerd. Bovendien hebben we de klinische effectiviteit en kosteneffectiviteit van een diagnostisch MRI-traject onderzocht bij mannen met een verdenking op PCa.

Hoofdstuk 2-4 geeft een overzicht van wat er in de huidige literatuur bekend is over de nauwkeurigheid van prostaat MRI. In **Hoofdstuk 2** beschrijven we de resultaten van een systematische literatuurstudie en meta-analyse naar de diagnostische accuratesse van mp-

MRI in het detecteren van alle vormen van PCa. Hierbij hebben we studies geïncludeerd die een MRI protocol gebruiken met T2-gewogen beeldvorming, gecombineerd met diffusie-gewogen beeldvorming en dynamische contrastversterkte MRI, zoals wordt aanbevolen door de European Society of Urogenital Radiology (ESUR). Zeven studies die voldeden aan de inclusiecriteria (526 patiënten) konden worden geanalyseerd. De gepoolde data laten een specificiteit zien van 0,88 (95% betrouwbaarheidsinterval (CI) 0,82-0,92) en een sensitiviteit van 0,74 (95% CI 0,66-0,81) voor de detectie van PCa, met negatief voorspellende waarden (NPV) van 0,65 tot 0,94. De methodologische kwaliteit van de geïncludeerde studies was redelijk, maar de studies bleken behoorlijk heterogeen te zijn. Desalniettemin lieten sensitiviteitsanalyses geen significante verschillen zien tussen de subgroepen. Op basis van de meta-analyse in dit hoofdstuk kunnen we concluderen dat er door de hoge specificiteit en redelijke sensitiviteit en NPV een rol lijkt te zijn voor mp-MRI bij de detectie van PCa.

In **Hoofdstuk 3** beschrijven we de diagnostische accuratesse van mp-MRI in het detecteren van klinisch significante PCa en analyseerden we de rol van Prostate Imaging Reporting and Data System (PI-RADS), een gestandaardiseerd systeem voor het beoordelen van een prostaat MRI. In deze diagnostische meta-analyse werden 14 studies (1785 patiënten) geanalyseerd. De gepoolde sensitiviteit was 0,78 (95% CI 0,70-0,84) en de gepoolde specificiteit 0,79 (95% CI 0,68-0,86) voor het detecteren van klinisch significant PCa, met NPV van 0,58 tot 0,95. Sensitiviteitsanalyses met studies waarbij PI-RADS correct wordt gebruikt, d.w.z. waarbij de methode duidelijk wordt omschreven en er niet wordt afgeweken van de aanbevelingen, lieten een gepoolde sensitiviteit zien van 0,82 (95% CI 0,72-0,89), met een specificiteit van 0,82 (95% CI 0,67-0,92). Studies die de PI-RADS aanbevelingen minder strikt gebruikten, een aangepaste versie gebruikten of de methode niet duidelijk verwoordden, lieten een gepoolde sensitiviteit zien van 0,73 (95% CI 0,62-0,82) en specificiteit van 0,75 (95% CI 0,61-0,84). Aanvullende sensitiviteitsanalyses bij studies die als primaire uitkomstmaat de detectie van klinisch significant PCa gebruikten, lieten een sensitiviteit van 0,84 (95% CI 0,76-0,89) en specificiteit van 0,75 (95% 0,66-0,83) zien. Bij studies die de detectie van alle vormen van PCa gebruikten, was de sensitiviteit 0,74 (95% CI 0,67-0,81) en de specificiteit 0,80 (95% CI 0,70-0,88). Op basis van de resultaten van de meta-analyse in dit hoofdstuk kunnen we concluderen dat de diagnostische accuratesse van mp-MRI voor de detectie van klinisch significant PCa redelijk goed is en deze lijkt te verbeteren wanneer PI-RADS correct wordt gebruikt. .

Hoofdstuk 4 beschrijft de resultaten van een systematische literatuurstudie naar de rol van MRI bij de lokale stadiëring van PCa. Meer specifiek was het doel om de diagnostische accuratesse van MRI voor de lokale stadiëring vast te stellen bij mannen met PCa welke is bewezen met prostaatbiopten. Bovendien was het doel om te analyseren wat hierbij de invloed is van verschillende MRI protocollen. Prostatectomie werd gebruikt als referentiestandaard. We hebben studies geïncludeerd die MRI gebruikten voor de detectie van extracapsulaire extensie (ECE; T3a), zaadblaasinvasie (SVI; T3b) of stadium T3 PCa. 75 studies (9796 patiënten) konden worden geanalyseerd. De gepoolde sensitiviteit en specificiteit voor ECE/T3a (45 studies, 5681 patiënten), SVI/T3b (34 studies, 5677 patiënten) en stadium T3 PCa (38 studies, 4001 patiënten) waren 0,57 (95% CI 0,49-0,64) en 0,91 (95% CI 0,88-0,93); 0,58 (95% CI 0,47-0,68) en 0,96 (95% CI 0,95-0,97); en 0,61 (95% CI 0,54-0,67) en 0,88 (95% CI 0,85-0,91), respectievelijk. Enkele van de onderzochte patiënt-, studie- en beeldvormingskarakteristieken hadden een duidelijke invloed op de nauwkeurigheid van de lokale stadiëring. Dit beïnvloedde met name de sensitiviteit, terwijl de specificiteit relatief stabiel bleef. Het gebruik van een MRI met een hogere veldsterkte (3,0 Tesla in plaats van 1,5 of 1,0 Tesla) en het gebruik van aanvullende multiparametrische MRI technieken leken de nauwkeurigheid te verbeteren. Het gebruik van een endorectale spoel liet geen additionele waarde zien voor ECE, maar verbeterde de sensitiviteit voor SVI wel enigszins.

De beslissing welk diagnostisch traject de voorkeur heeft, hangt naast de diagnostische accuratesse ook af van de kosteneffectiviteit. In Hoofdstuk 5 en 6 hebben we daarom ook de kosten en kosteneffectiviteit van het huidige diagnostisch traject (met TRUS-biopten) en het diagnostische MRI-traject bestudeerd. In **Hoofdstuk 5** hebben we de diagnostische meta-analyse uit Hoofdstuk 2 gebruikt als input voor een besliskundig model om zo de kosteneffectiviteit van de MRI-strategie (mp-MRI eventueel gevolgd door MRI-geleide biopten) te vergelijken met de huidige standaard (systematische TRUS-biopten). In dit model werden de kosten die gerelateerd waren aan de gebruikte technieken en de therapeutische consequenties van de tests meegenomen. Bovendien werden andere (in)directe consequenties, zoals kwaliteit van leven en overleving in het beslismodel verwerkt. Ondanks de onzekerheid rondom de kosteneffectiviteitschattingen, lijkt de MRI-strategie kosteneffectief voor het diagnosticeren van PCa ten opzichte van de huidige TRUS-strategie, wanneer een sensitiviteit van 20% of hoger wordt bereikt voor de MRI-geleide biopten. De hogere kosten van de MRI-strategie in het begin van het diagnostisch traject worden gecompenseerd doordat op langere termijn de behandelkosten af lijken te nemen, ten gevolge van minder vals positieve resultaten en de agressiviteit van de tumoren beter wordt ingeschat. De verbetering in kwaliteit van leven wordt bereikt doordat laag

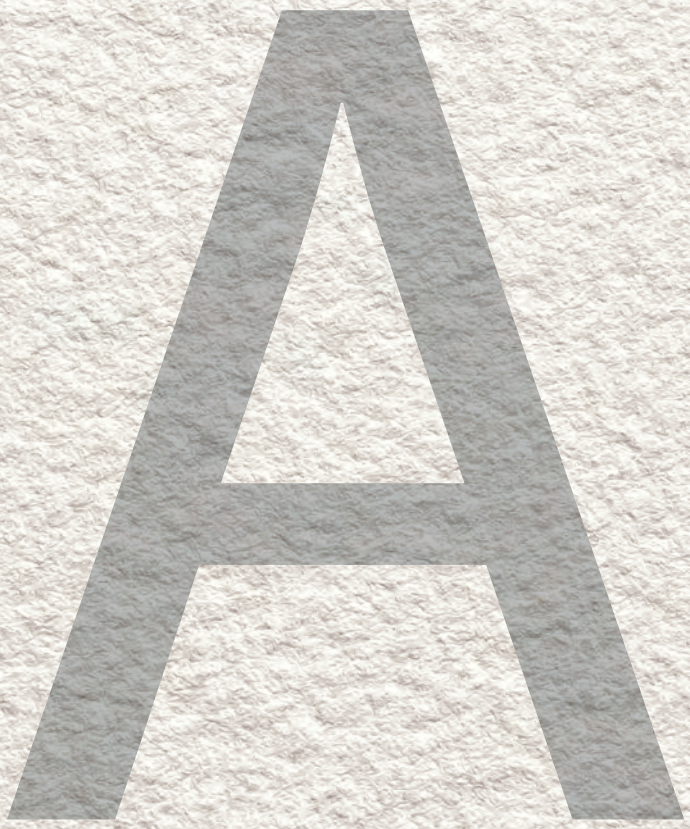
risico tumoren minder vaak onnodig invasief worden behandeld en doordat er minder klinisch significante tumoren gemist worden.

In **Hoofdstuk 6** geven we een inzicht in de gezondheidszorgkosten van urine-incontinentie (UI) die is ontstaan na PCa behandeling. Dit hebben we gedaan door gebruik te maken van een database van een zorgverzekeraar (Achmea Health Database; AHD), welke representatieve informatie bevat van circa 17% van de Nederlandse populatie. Het is belangrijk om de incidentie van UI betrouwbaar vast te stellen, omdat het een vaak voorkomende complicatie betreft die in veel gevallen irreversibel is. UI heeft grote implicaties voor de kwaliteit van leven, maar kan ook grote economische gevolgen hebben, aangezien de meeste mannen een gunstige overleving hebben. Door gebruik te maken van de AHD hadden we een unieke mogelijkheid de incidentie van UI en de bijbehorende kosten te vergelijken voor de verschillende behandelingsopties. De incidentie van UI was het hoogst na prostatectomie, waarbij de incidentie van UI hoger was na laparoscopische prostatectomie dan na een open prostatectomie. De hoge incidentie en de hoge kosten van UI na behandeling van PCa laten de omvang van het probleem zien. In de toekomst zullen deze kosten waarschijnlijk nog verder toenemen, tenzij nieuwe strategieën worden geïmplementeerd om de overdiagnose en overbehandeling van PCa te verbeteren.

Om de rol van een diagnostisch MRI-traject verder te onderzoeken, hebben we ook een prospectieve studie uitgevoerd, die wordt beschreven in **Hoofdstuk 7**. In deze studie hebben we de detectie van laag-, intermediair- en hoog-risico PCa met de MRI-strategie vergeleken met de huidige standaardzorg (TRUS-biopten). We hebben hiervoor 223 mannen geïncludeerd die nooit eerder prostaatbiopten hebben ondergaan. Deze 223 mannen hebben in het kader van de studie allemaal zowel een mp-MRI als systematische TRUS-biopten ondergaan. Mannen met afwijkende laesies op de mp-MRI hebben naast de systematische TRUS-biopten ook gerichte MRI-geleide biopten ondergaan. De resultaten van deze studie laten duidelijk de voordelen van een diagnostisch MRI-traject zien; het aantal mannen dat wordt gediagnosticeerd met een laag-risico PCa neemt af met 89,4%, terwijl het aantal mannen dat prostaatbiopten moet ondergaan afneemt met 51%. Bovendien neemt het aantal mannen waarbij een intermediair- of hoog-risico PCa wordt gevonden toe met 17,7%.

Ook al hebben de patiënt en zijn uroloog de hoofdrol in het diagnostische PCa-traject, toch zijn er veel andere betrokkenen rondom de zorg voor mannen met (een verdenking op) PCa; de familie van de patiënt, de huisarts, andere medisch specialisten en beleidsmakers. In

de **discussie** worden de belangrijkste bevindingen van dit proefschift besproken en geven we een overzicht van de belangrijkste uitdagingen voor deze verschillende betrokkenen. Bovendien geven we aanbevelingen met betrekking tot de mogelijke implementatie van mp-MRI.



Dankwoord (acknowledgements)

List of publications

Curriculum Vitae

DANKWOORD

Aan de voorkant van dit boekje staat maar één naam, maar ik heb dit proefschrift natuurlijk nooit kunnen volbrengen zonder steun, vertrouwen, inzet en gezelligheid van heel veel anderen. Iedereen die hier direct of indirect aan heeft bijgedragen, ben ik enorm dankbaar.
In de eerste plaats wil ik alle patiënten bedanken die deel hebben genomen aan de 'Brisbane studie'. Jullie bereidheid om mee te werken aan deze studie heeft er mede voor gezorgd dat in de toekomst ook andere mannen met een verdenking op prostaatkanker kunnen profiteren van verbeterde prostaatkankerdiagnostiek.

Vervolgens uiteraard mijn promotoren: prof. dr. M.M. Rovers en prof. dr. J.O. Barentsz. Wat heb ik het getroffen met jullie als begeleiders. Dankzij jullie vertrouwen in mij en de ruimte en mogelijkheden die jullie me hebben geboden gedurende mijn promotietraject, heb ik me zowel op wetenschappelijk/epidemiologisch vlak als op persoonlijk vlak verder kunnen ontwikkelen. Veel dank daarvoor.
Maroeska, buiten het feit dat je razendsnel, effectief en op de goede manier kritisch bent in het beoordelen van mijn stukken en opzetten van studies, is het bijzondere aan jou dat je daarnaast ook oog heb voor 'de persoon'. Je hebt me altijd weten te motiveren om het onderste uit de kan te halen en was flexibel op de momenten dat het nodig was. Mede dankzij de door jou geïnitieerde zomer BBQ's, groepsuitjes en borrels is de sfeer binnen de Evidence-Based Surgery-groep erg goed; daar kun je trots op zijn.
Jelle, jouw aanstekende enthousiasme en grenzeloze doorzettingsvermogen zijn enorm inspirerend. Je weet als geen ander hoe je je onderzoek moet profileren en dat altijd met één overkoepelend doel: de best mogelijke zorg voor de patiënt. Het is een voorrecht om hierbij betrokken te zijn en van te mogen leren.

Beste Janneke, je hebt het vermogen om klinische situaties die aanvankelijk ingewikkeld lijken op een heldere manier in een besliskundig model te integreren én dit vervolgens op een begrijpelijke manier te verwoorden. Het kosten-effectiviteitsstuk is daarvan het resultaat. Je hebt me veel kennis bijgebracht over health technology assessment en bent altijd bereid om mee te denken en advies te geven, waarvoor dank.

Beste prof. dr. J.A. Witjes, beste Fred. Bedankt dat je als mede-auteur bij veel van mijn stukken betrokken bent geweest. Jouw scherpe klinische blik heeft ervoor gezorgd dat in mijn stukken de connectie met de urologie altijd behouden is gebleven.

Beste Esther, bij sommige stukken ben jij eerste auteur en ik tweede, bij andere stukken is het andersom. Kortom, we hebben veel samengewerkt, met name voor het schrijven van systematic reviews en meta-analyses. Ook al hebben we eindeloze lijsten studies doorgespit voor kwaliteitsbeoordelingen en heb je soms toch nog even 'een paar' extra studies moeten scoren, toch bleef de samenwerking met jou altijd prettig. Bedankt voor je inzet, gezelligheid en de mooie studies die we samen hebben gepubliceerd.

I want to thank the research group of the Wesley Hospital in Brisbane (Australia), in particular Les Thompson and Morgan Pokorny, for the opportunity to be involved in their prospective study. Of course, I would also like to thank all the patients that were willing to participate; without you this important study was not possible.

Tijdens mijn promotietraject heb ik de mogelijkheid gehad om samen te werken met onderzoekers van het Achmea Kennisinstituut. Ik wil in het bijzonder Wieteke, Rogier, Han en Bianca bedanken voor de samenwerking. Het was heel leerzaam om vanuit jullie gezichtspunt naar (prostaatkanker)zorg te kijken.

Prof. dr. P.F.A. Mulders, prof. dr. M.A.A.J. van den Bosch en prof. dr. M.G.M. Hunink, beste leden van de manuscriptcommissie, bedankt dat jullie bereid zijn om zitting te nemen in de commissie en bedankt voor de snelle beoordeling van mijn manuscript.

Mijn eerste kennismaking met prostaatonderzoek was tijdens een extra curriculaire stage die ik heb gelopen voorafgaand aan mijn laatste co-schap. Dankzij deze stage, onder begeleiding van dr. ir. Tom Scheenen en dr. Thiele Kobus ben ik enthousiast geworden over prostaatonderzoek en kreeg ik de kans om dit een vervolg te geven na het behalen van mijn artsexamen. Tom en Thiele, bedankt voor jullie begeleiding en de mogelijkheden die jullie me hebben geboden.

Gedurende mijn promotietraject heb ik verschillende studenten begeleid tijdens hun (Bachelor)stage. Ik wil in het bijzonder Simone, Janine en Elsemieke bedanken voor hun inzet, enthousiasme en de gezelligheid.

Beste heer van den Berg, als voorzitter van de Prostaatkankerstichting en ervaringsdeskundige weet u als geen ander wat er belangrijk is voor mannen met (een verdenking op) prostaatkanker. Het was inspirerend om met u te spreken en meer te leren over het standpunt van 'de patiënt'. Bedankt voor uw openhartigheid.

Alle prostaatonderzoekers wil ik bedanken voor de collegialiteit en gezelligheid zowel op als buiten het werk. In het bijzonder de mede-biopteurs Bart, Caroline, Eline, Esther, Joyce, Kristian, Martijn H, Martijn S, Martin en Wulphert; bedankt voor jullie samenwerking en hulp bij het (leren) biopteren. Jurgen bedankt voor je mede-auteurschap bij mijn eerste meta-analyse. Bedankt ook voor alles wat ik van jou en de andere prostaatradiologen Marloes, Sjoerd, Patrik en Roel heb mogen leren op het gebied van prostaatkankerdiagnostiek. Marcel en Maarten bedankt voor de gezelligheid tijdens de koffiemomenten en congressen. Alle onderzoekers van de spectroscopie-groep, en in het bijzonder mijn oud-kamergenoten Marnix en Miriam, bedankt voor jullie warme welkom en gezelligheid tijdens mijn eerste 'stage' op de afdeling Radiologie.

Mijn kamergenoten bij de afdeling Operatiekamers en dan in het bijzonder Gerjon, Jan-Maarten, Tim, Tom, Casper en Jorte, wil ik bedanken voor al het enthousiasme, jullie hulp en de gezelligheid, ook buiten het werk. Ik heb veel van jullie geleerd en het is een fijn gevoel om mensen om je heen te hebben waar je op kunt rekenen. Door jullie voelde ik me altijd thuis op onze kamer en daar hebben de collectie koffiezetapparaten, worstenwoensdagen, visvrijdagen en borrels uiteraard aan bijgedragen. Ik weet zeker dat we nog met elkaar zullen blijven optrekken en/of samenwerken. Ook alle andere onderzoekers en stagaires van de Evidence-Based Surgery-groep wil ik bedanken voor de leuke en leerzame tijd samen.

Graag wil ik de secretariële ondersteuning en in het bijzonder Solange en Leonie van de afdeling Radiologie/Nucleaire geneeskunde en Alexandra en Miranda van de afdeling Operatiekamers bedanken voor alle hulp en regeldingen tijdens mijn promotietraject.

De leden van de Raad van Bestuur en het HR van het Radboudumc wil ik bedanken voor het opzetten van de Radboud Da Vinci Challenge. Het heeft me een rijker mens gemaakt en me in staat gesteld om me zowel op persoonlijk als op professioneel gebied te kunnen ontwikkelen. Jorik, Jimmy, Chella, Sandra, José, Janna, Jeannette, Mariska, Marcia, Jennifer, Kees en Ellen het was geweldig om deze 'reis' samen met jullie te beleven; ik denk er nog vrijwel dagelijks aan terug.

Mijn opleiders dr. Liesbeth Peters-Bax, dr. Marcel Janssen en prof. dr. L.J. Schultze Kool wil ik bedanken voor de mogelijkheid en ruimte om binnen de opleiding mijn promotie af te kunnen ronden. Ditzelfde geldt natuurlijk ook voor mijn groep collega AIOS in het Radboudumc. Ik zie ernaar uit om nog enkele jaren samen met jullie de opleiding te vervolgen en er een leerzame en gezellige tijd van te maken.

Alle (schoon)familie en mijn vrienden uit zowel Nijmegen als Tilburg wil ik bedanken voor de interesse in mijn werk en voor het plezier dat ik met jullie in de vrije tijd beleef. Ondanks dat de afstand en andere verplichtingen het ons soms lastig maken, hoop ik dat we elkaar vaak blijven zien.

Stijn en Frank, we zitten alle drie in ongeveer hetzelfde schuitje en weten daardoor exact hoe het is om te promoveren, AIOS te zijn en buiten het werk alle balletjes hoog te houden. Niet altijd makkelijk, maar met jullie als paranimfen en vrienden naast mijn zijde voel ik me altijd gesteund. Op naar de laatste verdediging!

Lieve ouders, dat ik hier nu sta, heb ik voor een groot deel aan jullie te danken. Ik kan me geen betere vader en moeder wensen. Bedankt dat jullie mij de vrijheid en mogelijkheden hebben gegeven om te ontdekken wat en wie ik wil worden en me hierin altijd onvoorwaardelijk hebben gesteund.

Lieve Claudia, toen ik startte met mijn promotietraject kende ik je pas 2 jaar. In de afgelopen jaren hebben we veel beleefd; we zijn samen gaan wonen in Malden, hebben 2 lieve kinderen gekregen en doen allebei het werk waar we mee verder willen. Mede dankzij jouw steun en geduld is het me gelukt om naast vader en echtgenoot ook doctor te worden. Het is fijn om bij elkaar te zijn en ik zie uit naar onze toekomst samen!

LIST OF PUBLICATIONS

Kobus T, Vos PC, Hambrock T, **de Rooij M**, Hulsbergen-Van de Kaa CA, Barentsz JO, Heerschap A, Scheenen TW. Prostate cancer aggressiveness: in vivo assessment of MR spectroscopy and diffusion-weighted imaging at 3 T. *Radiology* 2012; 265(2): 457-67.

Fortuin A, **de Rooij M**, Zamecnik P, Haberkorn U, Barentsz JO. Molecular and functional imaging for detection of lymph node metastases in prostate cancer. *International journal of molecular sciences* 2013; 14(7): 13842-75.

de Rooij M, Hamoen EH, Futterer JJ, Barentsz JO, Rovers MM. Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis. *AJR American journal of roentgenology* 2014; 202(2): 343-51.

de Rooij M, Crienen S, Witjes JA, Barentsz JO, Rovers MM, Grutters JP. Cost-effectiveness of magnetic resonance (MR) imaging and MR-guided targeted biopsy versus systematic transrectal ultrasound-guided biopsy in diagnosing prostate cancer: a modelling study from a health care perspective. *European urology* 2014; 66(3): 430-6.

de Rooij M, Crienen S, Witjes JA, Barentsz JO, Rovers MM, Grutters JP. Reply to Sarah Willis, Alec Miners, and Jan van der Meulen's letter to the editor re: Maarten de Rooij, Simone Crienen, J. Alfred Witjes, Jelle O. Barentsz, Maroeska M. Rovers, Janneke P.C. Grutters. Cost-effectiveness of magnetic resonance (MR) imaging and MR-guided targeted biopsy versus systematic transrectal ultrasound-guided biopsy in diagnosing prostate cancer: a modelling study from a health care perspective. *European urology* 2014; 66(2): e30.

Pokorny MR, **de Rooij M**, Duncan E, et al. Prospective study of diagnostic accuracy comparing prostate cancer detection by transrectal ultrasound-guided biopsy versus magnetic resonance (MR) imaging with subsequent MR-guided biopsy in men without previous prostate biopsies. *European urology* 2014; 66(1): 22-9.

Hamoen EH, **de Rooij M**, Witjes JA, Barentsz JO, Rovers MM. Measuring health-related quality of life in men with prostate cancer: A systematic review of the most used questionnaires and their validity. *Urologic oncology* 2015; 33(2): 69 e19-28.

Hamoen EH, **de Rooij M**, Witjes JA, Barentsz JO, Rovers MM. Use of the Prostate Imaging Reporting and Data System (PI-RADS) for Prostate Cancer Detection with Multiparametric Magnetic Resonance Imaging: A Diagnostic Meta-analysis. *European urology* 2015; 67(6): 1112-21.

de Rooij M, Hamoen EH, Witjes JA, Barentsz JO, Rovers MM. Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis. *European urology* 2016; 70(2): 233-45.

de Rooij M, Nieuwboer W, Smit RNG, Witjes JA, Barentsz JO, Rovers MM. The economic burden of urinary incontinence after prostate cancer treatment. *Submitted*.

Venderink W, **de Rooij M**, Sedelaar JP, Huisman HJ, Futterer JJ. Elastic Versus Rigid Image Resitration in Magnetic Resonance Imaging–transrectal Ultrasound Fusion Prostate Biopsy: A Systematic Review and Meta-analysis. *European urology focus* 2016.

Venderink W, Govers TM, **de Rooij M**, Futterer JJ, Sedelaar JP. Cost-Effectiveness Comparison of Imaging-Guided Prostate Biopsy Techniques: Systematic Transrectal Ultrasound, Direct In-Bore MRI, and Image Fusion. *AJR American journal of roentgenology* 2017: 1-6.

Barentsz JO, **de Rooij M**, Villeirs G, Weinreb J. Prostate Imaging-Reporting and Data System Version 2 and the Implementation of High-quality Prostate Magnetic Resonance Imaging. *European urology* 2017.

CURRICULUM VITAE

Maarten de Rooij was born in Tilburg on October 10, 1983. After graduating from secondary school (Odulphus Lyceum, Tilburg) he started with medical school at the Radboud University Nijmegen in 2004. After attending several electives he became interested in Radiology.

To explore both the scientific and clinical aspects of Radiology, he did an extracurricular internship at the department of Radiology of the Radboud university medical center. Under supervision of dr. ir. Tom Scheenen and dr. Thiele Kobus he became familiar with prostate cancer research, which is one of the main scientific research themes of the hospital.

He obtained his medical degree in September 2011 and subsequently started his PhD under supervision of prof. dr. Maroeska Rovers (Professor of Evidence-Based Surgery) and prof. dr. Jelle Barentsz (Professor of Radiology). During his PhD he also gained clinical experience by performing MR guided in-bore prostate biopsies, attended several Epidemiology Master courses to obtain a registration as epidemiologist, and supervised Bachelor students. Furthermore, he participated in the Radboud Da Vinci Challenge, a one-year program that offers excellent PhD candidates the opportunity to experience broad personal development.

In 2014, Maarten was awarded the Lauterbur Award from the Society of Computed Body Tomography and Magnetic Resonance for part of his research.

Since January 2015, Maarten is a resident in Radiology at the Radboud university medical center. He lives together with Claudia van IJzendoorn and their two children Emma and Philip.

CURRICULUM VITAE

Maarten de Rooij werd geboren op 10 oktober 1983 in Tilburg. Na het behalen van zijn VWO diploma aan het Odulphus Lyceum te Tilburg, startte hij in 2004 met zijn studie geneeskunde aan de Radboud Universiteit Nijmegen.

Tijdens de studie groeide zijn interesse voor het vak Radiologie, mede door het volgen van enkele keuzevakken. Dit gaf hij verder vorm door tijdens de wachttijd voor zijn laatste coschap in Tanzania een extracurriculaire stage bij de afdeling Radiologie in het Radboudumc te volgen. Gedurende deze stage naar de rol van MRI bij het bepalen van agressiviteit van prostaatkanker, onder begeleiding van dr. ir. Tom Scheenen en dr. Thiele Kobus, ontstond de interesse in zowel de klinische als de wetenschappelijke kant van het vak.

Na het behalen van zijn artsexamen startte hij in september 2011, onder supervisie van prof. dr. Maroeska Rovers (hoogleraar Evidence-Based Surgery) en professor Jelle Barentsz (hoogleraar Radiologie), aan zijn promotietraject. Tevens deed hij klinische ervaring op met het verrichten van MRI-geleide prostaatbiopten, volgde hij vakken van de Master Epidemiologie voor het behalen van zijn registratie tot epidemioloog en begeleidde hij verschillende Bachelor studenten. Ook nam hij deel aan de Radboud Da Vinci Challenge, een programma voor excellente promovendi gericht op een brede persoonlijke ontwikkeling.

In 2014 won Maarten met een presentatie van een deel van zijn onderzoek de Lauterbur Award van de Society of Computed Body Tomography and Magnetic Resonance.

Sinds januari 2015 is Maarten als radioloog in opleiding werkzaam in het Radboudumc. Maarten woont samen met Claudia van IJzendoorn en hun twee kinderen Emma en Philip.