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

G Mowatt,[1]* G Scotland,[1,2] C Boachie,[1] M Cruickshank,[1] JA Ford,[1] C Fraser,[1] L Kurban,[3] TB Lam,[4] AR Padhani,[5] J Royle,[3] TW Scheenen[6] and E Tassie[2]

[1]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[2]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK
[3]Aberdeen Royal Infirmary, Aberdeen, UK
[4]Academic Urology Unit, University of Aberdeen, Aberdeen, UK
[5]Mount Vernon Cancer Centre, Northwood, Middlesex, UK
[6]Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands

*Corresponding author

Executive summary

Localisation of prostate abnormalities for biopsy

Health Technology Assessment 2013; Vol. 17: No. 20
DOI: 10.3310/hta17200

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Executive summary

Background

In the UK, prostate cancer (PC) is the most common cancer in men. Many men find themselves with the dilemma of having an elevated prostate-specific antigen (PSA) level and a negative prostate biopsy, and the best way for doctors to manage these patients remains uncertain. The strategy of further repeat biopsies for these men remains controversial, with uncertainties surrounding the optimal number of cores, which area of the prostate to target, and imaging modality for guidance. This has led to the introduction of new imaging techniques. Conventional standard (T2-weighted) magnetic resonance imaging (T2-MRI) can be performed with add-on modalities, including three-dimensional magnetic resonance spectroscopy (MRS), dynamic contrast-enhanced MRI (DCE-MRI) and diffusion-weighted MRI (DW-MRI).

Objectives

This review aims to assess the diagnostic accuracy of MRS and enhanced MRI techniques (DCE-MRI, DW-MRI) and the clinical effectiveness and cost-effectiveness of strategies involving their use in aiding the localisation of prostate abnormalities for biopsy in patients with prior negative biopsy in whom there remains a clinical suspicion that they are harbouring malignancy.

Methods

Electronic databases searched included MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, Bioscience Information Service (BIOSIS), Science Citation Index (SCI), Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE), Medion, Health Technology Assessment database, conference abstracts from the American Society for Clinical Oncology (ASCO) and current research registers. Searches were carried out from 1995 to March 2012. Types of studies considered were direct studies or randomised controlled trials reporting absolute numbers of true- and false-positives and true- and false-negatives, allowing the calculation of sensitivity, specificity or predictive values. The population was men with suspected PC and elevated PSA level but previously negative biopsy. Index tests were MRS, DCE-MRI and DW-MRI, and comparator tests were standard T2-MRI and transrectal ultrasonography (TRUS). The reference standard was histopathological assessment of biopsied tissue obtained via transrectal needle biopsy, saturation biopsy, transperineal template biopsy or from prostatectomy specimens.

Two reviewers independently screened the titles and abstracts of all reports identified by the search strategy and full-text papers were subsequently obtained for assessment. Data extraction was undertaken by one reviewer and checked by a second. Two reviewers independently assessed the risk of bias of the diagnostic studies using a modified version of the QUADAS-2 (quality assessment of diagnostic accuracy studies, version 2) instrument.

The results of the individual studies were tabulated and sensitivity, specificity and their 95% confidence intervals (CIs) presented for each test or combination of tests at both patient and biopsy level. The presence of heterogeneity was assessed by visual examination of pairs of forest plots of sensitivity and specificity. Separate summary receiver operating characteristic (SROC) curves were derived for different levels of analysis. Meta-analysis models were fitted using hierarchical SROC (HSROC) curves. Summary sensitivity, specificity, positive and negative likelihood ratios and diagnostic odds ratios for each model were reported as median and 95% CI. An indirect comparison of tests was also undertaken.

An economic model was developed to assess the cost-effectiveness of using alternative MRS/MRI sequences to direct TRUS-guided biopsies (TRUS/Bx), compared with the standard practice of relying on systematic extended TRUS-guided biopsies (in patients with a previous negative biopsy). The alternative diagnostic pathways were embedded in a Markov model simulating the progression of undiagnosed cancer and the downstream impact of diagnosis and treatment on survival and health-related quality of life (QoL). Costs incorporated in the model included the costs associated with obtaining the final diagnosis (cancer/no cancer), management of biopsy complications, cancer staging, cancer treatment, and the management of complications resulting from cancer treatment. Survival benefits of diagnosis were captured through the application of relative risk parameters reflecting the benefit of appropriate treatment by stage of underlying cancer. Health-state utilities associated with cancer stage and the occurrence of treatment complications were incorporated in the model to estimate quality-adjusted life-years (QALYs). Experimental strategies were compared incrementally with standard practice in terms of their incremental cost per life-year and QALY gained.

Results

Number and quality of studies

Fifty-one studies (39 full text and 12 abstracts) were included, involving over 10,000 men. Only full-text studies were assessed for risk of bias, the majority of which were considered to have a low risk of bias for the patient selection (74%, 29/39), index test (100%, 39/39) and flow and timing (92%, 36/39) domains. In the reference standard domain, the majority of studies (64%, 25/39) were considered at high risk of bias owing to a lack of follow-up.

Summary of benefits and risks

In meta-analyses of the individual tests, sensitivity was highest for MRS at 92% (95% CI 86% to 95%), followed by T2-MRI at 86% (95% CI 74% to 93%) and DCE-MRI at 79% (95% CI 69% to 87%), whereas specificity was highest for TRUS (used as an imaging test) at 81% (95% CI 77% to 85%), followed by MRS at 76% (95% CI 61% to 87%). In pooled estimates for combinations of tests, sensitivity was highest for 'MRS or T2-MRI' at 96% (95% CI 90% to 98%) followed by 'DCE-MRI or T2-MRI' at 88% (95% CI 80% to 96%), whereas specificity was highest for 'MRS and T2-MRI' at 74% (95% CI 65% to 84%). Only one small study involving 43 patients reported DW-MRI, with sensitivity of 100% (specificity not reported). The results of the indirect comparison broadly reflected those of the meta-analyses of the individual tests and combinations of tests.

Summary of costs

The base-case analysis showed average discounted lifetime costs to range between £3895 using systematic TRUS-guided biopsies and £4056 using positive findings on either T2-MRI or DCE-MRI to determine and direct biopsies (60-year-old cohort, cancer prevalence 24%). The corresponding figures for the same strategies in a 70-year-old cohort were £3199–3660. Using T2-MRI to direct biopsies represented the least costly approach in low-prevalence (10%) cohorts.

Summary of cost-effectiveness

Survival and QALY differences between strategies were very small but these favoured more sensitive approaches. Under base-case parameter values and assumptions (with underlying cancer prevalence 24%), the incremental cost-effectiveness ratio (ICER) for T2-MRI was < £30,000 per QALY in comparison with systematic extended-cores TRUS/Bx (all cohorts) and T2-MRI was found to dominate extended-cores TRUS/Bx in low-prevalence cohorts. However, probabilistic sensitivity analysis demonstrated a high degree of uncertainty surrounding the incremental cost-effectiveness of T2-MRI compared with extended-cores TRUS/Bx in the moderate prevalence cohorts. The cost-effectiveness of MRS compared with T2-MRI was less favourable under base-case assumptions, although its ICER did fall to < £30,000 compared with extended-cores TRUS/Bx in the moderate prevalence 60-year-old cohort, and also compared with T2-MRI-directed

biopsy in the high-prevalence 60-year-old cohort. The ICER for MRS, or any of the other more sensitive strategies, did not fall to < £30,000 in any of the 70-year-old cohorts under base-case assumptions.

Sensitivity analyses

Base-case findings were found to be highly sensitive to a number of uncertain parameters and assumptions. The cost-effectiveness of using MRS to direct biopsies was found to be particularly sensitive to the cost of prostate biopsies relative to the cost of obtaining a MRS sequence. When the cost of obtaining biopsies was raised by ~£115 relative to the cost of MRS, MRS-directed biopsy was found to dominate extended-cores TRUS/Bx in all of the cohorts, and its ICER dropped to < £30,000 in comparison with the T2-MRI-directed approach in the moderate- and high-prevalence 60-year-old cohorts (although it remained > £30,000 in all of the 70-year-old cohorts). The cost-effectiveness of MRS was also crucially sensitive to its modelled ability to discriminate between low- and moderate-/high-risk cancer. When all of its false-negative findings were modelled to occur in patients with low-risk disease, its cost-effectiveness improved substantially in the moderate- and high-prevalence 60-year-old cohorts, although its ICERs remained less favourable in the 70-year-old cohorts. Factors undermining the cost-effectiveness of MRS included the application of lower disease progression rates and lower relative risk reductions associated with diagnosis and treatment. Although a lack of available evidence precluded its inclusion in our base-case analysis, if DW-MRI could be shown to perform similarly to MRS in terms of diagnostic accuracy, it would probably be favoured over MRS for its lower cost.

Discussion

Strengths, limitations of the analyses and uncertainties

In terms of strengths, a comprehensive literature search was undertaken. A HSROC model was used, which takes account of the trade-off between true/false-positives and models between-study heterogeneity. Pooled estimates were performed at both patient and biopsy level and an indirect comparison of tests was undertaken. In terms of limitations, non-English-language studies were excluded. Few studies reported DCE-MRI or DW-MRI or included a period of follow-up as part of the reference standard. The index and comparator tests were not independent of the reference standard.

In terms of uncertainties, where studies reported an 'equivocal' results category, this was classed with positive rather than negative results, increasing sensitivity and decreasing specificity, whereas the reverse would have been the case if 'equivocal' had been classed with negative results. There was only limited evidence available of the ability of MRS and other MRI techniques to detect clinically significant disease. In studies reporting MRS or other MRI techniques a systematic TRUS/Bx was also undertaken and in most of these studies it was unclear how this contributed to sensitivity and specificity values reported.

Generalisability of the findings

All studies included in the pooled estimates reported men with suspected PC and elevated PSA level but previously negative biopsy, and therefore these findings would be broadly generalisable to patients meeting the above criteria. However, in one study the spectrum of patients was not representative (all had atypical small acinar proliferation). In two studies imaging was MR-guided (rather than TRUS-guided), a method not generally used in the UK. Six studies reporting TRUS-guided systematic biopsies were large screening studies, which is not representative of how men are detected with PC in the UK.

Conclusions

Implications for service provision

Given the level of uncertainty surrounding several key model inputs, it is difficult to arrive at definitive conclusions on the cost-effectiveness of using different MRS/MRI sequences to aid the localisation of prostate abnormalities for biopsy. However, our modelling suggests that, under certain circumstances,

T2-MRI may be considered cost-effective in comparison with systematic TRUS/Bx, and if MRS and DW-MRI can be shown to have high sensitivity for detecting moderate-/high-risk cancer, while negating the need for patients with no cancer or low-risk disease to undergo biopsy, then their use could represent a cost-effective approach to diagnosis.

The introduction of MRS and other MRI techniques (T2-MRI, DCE-MRI, DW-MRI) for evaluation of men with negative TRUS/Bx but in whom there remains suspicion of cancer would have a range of implications for the NHS. These would arise primarily because of a shift in the test–treatment pathway for this group, with changes in the method of making diagnosis resulting in changes to the types of patients being treated, offered patient options and timings of treatments. This would have consequential effects on service provision, costs and training. If urological and/or radiological services were to undertake targeted biopsies of MRS-/MRI-suspicious regions then extra provision would be required for this. A new generation of equipment and software would be needed to enable accurate, documentable biopsies to be obtained from all regions of the prostate. If MRS/MRI identified more patients with localised disease with intermediate and high risk of progression then this would increase the proportion of patients considered eligible for radical therapies. If MRS or MRI detected few patients with low risk of disease progression then fewer patients in this category would undergo perhaps inappropriate radical therapies. Thus, the total number of patients undergoing radical therapies would be appropriately decreased, requiring a rebalancing of resources currently allocated to surgical and radiation therapy services. Furthermore, if MRS or MRI contributed to the more accurate classification of patients with a low risk of progression, this would lead to an increase in the proportion of appropriately selected patients who are likely to undergo 'active surveillance', helping to mitigate the current high dropout rate of this approach. The implications for the follow-up of active surveillance patients would include repeated PSA testing, repeated interval biopsies and follow-up clinics (much of this work is protocol driven and could be nurse practitioner led). Taken together, earlier, more accurate diagnoses and more appropriate treatments of PC may improve patient outcomes by reducing treatment-related morbidity, improving survival and, in the longer term, reducing the requirement for end-of-life and palliative care services. There would be cost implications of these service reconfigurations and for changes in treatment patterns mentioned above. Implementation would also result in the need for further training of all staff involved in delivering care to patients with PC.

Suggested research priorities

Prospective studies are required in men with suspected PC in whom PSA level is elevated but a previous biopsy has been negative, comparing the utility of the individual and combined components of a multiparametric magnetic resonance (MR) approach (MRS, DCE-MRI and DW-MRI) with both a MR-guided or -directed biopsy session and an extended 14-core TRUS/Bx scheme against a reference standard of histopathological assessment of biopsied tissue obtained via saturation biopsy, template biopsy or prostatectomy specimens. A follow-up time of 12 months should form part of the reference standard. Investigations of DW-MRI should be encouraged, as it is already gaining widespread acceptance in the clinic owing to its relatively easy use. These studies should also report the sensitivity of the tests in detecting clinically significant disease (Gleason score of ≥ 7 and/or volume > 0.5 ml). In addition to diagnostic outcomes, adverse event data and impact of the tests on subsequent physician attitudes to patient management should also be obtained, as well as cost-effectiveness data including impact of testing on health-related QoL.

Uncertainties surrounding cost-effectiveness could be significantly reduced by future research focusing on generating comparable estimates of (1) the sensitivity of MRI-/MRS-directed and systematic approaches to TRUS/Bx (using a robust and common reference standard); (2) the prospective sensitivity or specificity of MRS or MRI sequences for detecting different grades of localised disease in the repeat biopsy setting; and (3) the full economic costs of MRI sequences and systematic approaches to TRUS/Bx based on different numbers of cores.

Further, with the survival and QALY differences between strategies being so small, and of questionable clinical significance, the choice between strategies might be better informed by patient or public

preferences for process of care factors to which the standard QALY model may be insensitive. Scope exists to carry out preference elicitation studies to identify and value the key factors influencing patients' preferences for alternative diagnostic, monitoring, and subsequent treatment pathways.

Study registration

This study is registered as PROSPERO CRD42011001376.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Publication

Mowatt G, Scotland G, Boachie C, Cruickshank M, Ford JA, Fraser C, *et al*. The diagnostic accuracy and cost-effectiveness of magnetic resonance spectroscopy and enhanced magnetic resonance imaging techniques in aiding the localisation of prostate abnormalities for biopsy: a systematic review and economic evaluation. *Health Technol Assess* 2013;**17**(20).

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.596

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/146/01. The contractual start date was in August 2010. The draft report began editorial review in July 2012 and was accepted for publication in January 2013. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Mowatt *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.