

MRI software and cognitive fusion biopsies in people with suspected prostate cancer: a systematic review, network meta-analysis and cost-effectiveness analysis

Alexis Llewellyn,¹ Thai Han Phung,² Marta O Soares,²
Lucy Shepherd,¹ David Glynn,² Melissa Harden,¹
Ruth Walker,¹ Ana Duarte² and Sofia Dias^{1*}

¹Centre for Reviews and Dissemination, University of York, York, UK

²Centre for Health Economics, University of York, York, UK

*Corresponding author sofia.dias@york.ac.uk

Published October 2024

DOI: 10.3310/PLFG4210

Scientific summary

MRI software and cognitive fusion biopsies in people with suspected prostate cancer: a systematic review, network meta-analysis and cost-effectiveness analysis

Health Technology Assessment 2024; Vol. 28: No. 61

DOI: 10.3310/PLFG4210

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

Scientific summary

Background

Prostate cancer (PCa) is the most commonly diagnosed cancer in men in the UK. In the NHS people with suspected PCa are offered multiparametric magnetic resonance imaging (mpMRI). People with suspected PCa, according to MRI, are offered a biopsy procedure to confirm the presence and severity of cancer. Traditionally patients were offered a systematic transrectal, ultrasound-guided prostate biopsy (or systematic biopsy). Since the introduction of mpMRI, specific areas of abnormal tissue can be targeted, by combining (or fusing) the results of mpMRI and ultrasound imaging. Several methods for fusing MRI and ultrasound images exist, including cognitive fusion (CF), in which a region of interest is identified prior to biopsy and the biopsy operator estimates where it might be on an ultrasound image, and software fusion (SF), where regions of interest on magnetic resonance images are identified and contoured before biopsy and overlaid with the prostate contours on ultrasound images during the biopsy. Systematic biopsy may be used in addition to targeted biopsy. A number of SF technologies are available. However, the effectiveness and cost-effectiveness of SF compared with CF is uncertain.

Objectives

This study aimed to assess the clinical and cost-effectiveness of SF biopsy systems in people with suspected localised and locally advanced PCa.

Methods

Systematic review

A systematic review of the diagnostic accuracy, clinical effectiveness, safety and practical implementation of nine SF systems compared with CF and with each other, in people suspected PCa according to MRI was conducted.

Comprehensive bibliographic searches, including MEDLINE and EMBASE and supplementary sources, were conducted up to 2 August 2022 for published and unpublished literature.

Studies of people with suspected PCa who have had a MRI scan that indicates a significant lesion [Likert or prostate imaging – reporting and data system (PI-RADS) score of 3 or more], including biopsy-naïve and repeat biopsy patients with a previous negative prostate biopsy, and comparing SF with CF or with another SF device, were included. The following SF technologies were included: ARTEMIS (InnoMedicus ARTEMIS), BioJet (Healthcare Supply Solutions Ltd), BiopSee (Medcom), bkFusion (BK Medical UK Ltd and MIM Software Inc.), Fusion Bx 2.0 (Focal Healthcare), FusionVu (Exact Imaging), iSR'obot Mona Lisa™ (Biobot iSR'obot), KOELIS Trinity (KOELIS and Kebomed) and UroNav Fusion Biopsy System (Phillips). Previous versions were also eligible. In-bore (or in-gantry) biopsies were excluded. Prospective, randomised and non-randomised comparative studies were included, and retrospective evidence where no prospective evidence could be found for an eligible SF device. To provide sufficient evidence for a network meta-analysis (NMA), within-patient comparisons or randomised controlled trials (RCTs) between SF and systematic biopsy, and between CF and systematic biopsy, were also eligible to inform indirect comparisons of diagnostic accuracy.

Two researchers independently screened the titles and abstracts of all reports identified by the bibliographic searches and of all full-text papers subsequently obtained. Data extraction and quality

assessment were conducted by at least one researcher and checked by a second. Risk of bias of diagnostic accuracy studies was assessed using quality assessment of diagnostic accuracy studies-comparative (QUADAS-C).

For diagnostic accuracy outcomes, studies reporting sufficient data were included in network meta-analyses comparing SF and CF with or without concomitant systematic biopsy, and systematic biopsy alone, where odds of being categorised in each of different cancer grades were allowed to vary by biopsy type. Results were reported as odds ratios with 95% credible intervals (CrIs). Additional diagnostic accuracy results that could not be pooled in a meta-analysis and clinical effectiveness, safety and implementation outcomes were synthesised narratively.

Economic analysis

Cost-effectiveness evidence comparing SF biopsy systems with CF for targeted prostate biopsy in men with suspected PCa was identified by the previously mentioned searches, with evidence narratively summarised and tabulated. Studies were appraised for their quality, generalisability and appropriateness to inform the decision problem as defined by the National Institute for Health and Care Excellence Diagnostics Assessment Report (NICE DAR) scope. A targeted search was conducted to identify evidence to support the development of a de novo decision model. The searches aimed to identify cost-effectiveness evidence of diagnostic strategies at the point of biopsy to support the model conceptualisation. Evidence was reviewed to (1) identify value components of the biopsy approaches, (2) characterise alternative mechanisms of evidence linkage from disease prevalence, diagnostic accuracy, choice of treatment to final outcomes, and (3) identify any UK-relevant sources of evidence. A de novo decision analytic model was developed to estimate the cost-effectiveness of SF compared to CF. The model evaluated two strategies for two alternative comparisons: (1) targeted SF biopsy versus targeted cognitive biopsy and (2) combined (targeted and systematic) SF biopsy versus combined cognitive biopsy. The four strategies could not be incrementally compared due to the mechanism of evidence generation for the diagnostic accuracy, which relied on separate evidence networks.

The de novo model consisted of two components: (1) a decision tree, which captured biopsy adverse events (AEs), repeated biopsies and classified individuals according to their biopsy results and underlying true disease status, and (2) long-term model to link classification to clinical management decisions and this to longer-term costs and consequences (e.g. disease progression and PCa mortality) so that differences in costs, life-year gains and quality-adjusted life-years (QALYs) were quantified over a lifetime horizon.

The model required the development of (1) an extension to the evidence synthesis to allow quantifying the extension of test misclassification in the diagnostic model with SF biopsy and CF biopsy, and (2) an inference model to derive unobservable transition probabilities for the long-term model.

Results

The systematic review of clinical evidence included a total of 3733 patients who received SF and 2154 individuals with CF from 23 studies. Evidence was included for all devices specified in the protocol, except for Fusion Bx 2.0 and FusionVu. Overall, the evidence for all devices was at high risk of bias. Overall, biopsy-naïve patients were under-represented. Fourteen studies were included in the meta-analyses.

Diagnostic accuracy

Across all analyses results must be interpreted with caution due to the high risk of bias in the evidence base and wide uncertainty over the results. The meta-analyses show that patients undergoing SF biopsy may have: (1) a lower probability of being classified as not having cancer, (2) similar probability of being classified as having non-clinically significant cancer [International Society of Urological Pathology (ISUP)

grade 1], and (3) higher probability of being classified at higher ISUP grades, particularly ISUP 2. Similar results were obtained where both biopsy methods were combined with systematic biopsy.

Additional meta-analyses of cancer detection rates suggest that, compared with CF biopsy, SF may identify more PCa (any grade) (OR 1.30; 95% CrI 1.06, 1.61). Adding systematic biopsy to cognitive or SF may increase the detection of all PCa and of clinically significant (CS) cancer, and from this evidence there is no suggestion that SF with concomitant systematic biopsy is superior to CF with systematic biopsy.

Meta-analyses of cancer detection rates, by individual device, showed that compared with CF biopsy, BioJet and Urostation are associated with a higher detection of PCa overall. There was no evidence that any of the SF devices increased detection of CS cancer (except for BioJet, although this is based on one low-quality study), and overall, the evidence was insufficient to conclude whether any individual devices were superior to CF, or whether some SF technologies are more accurate than others.

Clinical effectiveness

There is no evidence that biopsy positivity rates and safety outcomes differ significantly between SF and CF, or between SF devices. There was some evidence that systems with rigid registration (BioJet or UroNav) are easier and faster to use than elastic registration (KOELIS Trinity), although this is informed by a single, small study and is not conclusive.

Cost-effectiveness

One full cost-effectiveness study of SF compared targeted SF to targeted CF. However, the findings of the study were not considered generalisable to the decision problem under assessment. Sixteen studies were identified of which nine were selected to inform the conceptualisation and parameterisation of the de novo decision model.

The base-case cost-effectiveness analysis suggests for the targeted biopsy and the combined biopsy comparisons, that SF strategy is on average costlier and yields greater QALYs than the CF strategy, resulting in a probabilistic incremental cost-effectiveness ratio (ICER) of £6197 and £2199 per additional QALY for each comparison, respectively. These ICERs are below the lower bound of the cost-effectiveness threshold range recommended by NICE, suggesting that SF may be cost-effective compared to CFs in both the targeted and the combined comparisons. However, these results should be interpreted cautiously given the uncertainties in the relative diagnostic accuracy evidence which informs the model. The probabilistic analysis suggests a higher probability of cost-effectiveness for SF versus CF at the range of cost-effectiveness thresholds recommended by NICE (0.64 and 0.68 at £20,000 and £30,000 per additional QALY for targeted SF biopsy).

Discussion

This assessment includes a broad, comprehensive literature search for software and CF technologies and has been conducted following recognised guidelines to ensure high quality. The review identified evidence on the diagnostic accuracy of nine SF technologies, and is the first systematic review to formally compare the relative accuracy of SF and CF, with and without systematic biopsy, as well as different SF devices, using both direct and indirect evidence in a NMA. Unlike recent systematic review evidence, our review found that SF increased detection of clinically insignificant cancer compared with CF.

Our review has a number of limitations. The evidence included in the systematic review is at high risk of bias overall. There was variation in patient and study characteristics. Biopsy-naïve patients, who form the large majority of patients eligible for targeted biopsy, were under-represented, although there was insufficient evidence to evaluate whether the relative accuracy of software and CF differed between

biopsy-naïve and repeat biopsy patients. There was insufficient evidence to explore the impact of a number of other potential effect modifiers, including lesion location, operator experience, biopsy routes and anaesthesia methods. There were few studies per comparison, not all studies reported outcomes by all cancer grades, and most estimates from the meta-analyses were imprecise, particularly at higher cancer grades where data were most sparse. The network meta-analyses relied on the assumption that CF was equivalent across different centres, which is uncertain.

No evidence was found for most of this assessment's prespecified outcomes: biopsy sample suitability/quality, number of repeat biopsies performed, procedure completion rates, software failure rate, time to diagnosis, length of hospital stay, time taken for MR image preparation, subsequent PCa management, re-biopsy rate, hospitalisation, overall survival, progression-free survival (PFS), patient- and carer-reported outcomes [including tolerability and health-related quality of life (HRQoL)], barriers and facilitators to implementations.

The cost-effectiveness results are driven by the modelled differences in diagnostic accuracy between software and CF, particularly the increased correct detection of Cambridge Prognostic Group 1 (CPG 1) (resulting in net losses for SF) and CPG 2 (resulting in net gains for SF). The External Assessment Group (EAG)'s NMA and its extension underpinned the economic model, so its limitations apply to the cost-effectiveness estimates. The magnitude of value realised for SF, compared with CF, depends on the balance between different degrees of misclassification and correct classification with the two technologies and on the prevalence of disease at each cancer grade. The value of SF is thus driven by comparative diagnostic accuracy (compared to 'gold standard') derived where evidence is particularly sparse (cancer grades above 2), and by prevalence, which is also affected by evidence sparsity. Therefore, the estimates of cost-effectiveness are affected by unquantified uncertainty and should be interpreted with caution.

Conclusions

Compared to CF biopsy, patients undergoing SF biopsy may show a lower probability of being classified as not having cancer, similar probability of being classified as having non-CS cancer, and a higher probability of being classified at higher ISUPs, particularly ISUP 2. Both SF and CF biopsy can miss CS cancer lesions, and the addition of standard-systematic biopsy increases the detection of all PCa and CS cancer for both fusion methods. There is insufficient evidence to conclude on the relative accuracy and clinical effectiveness of different software devices.

Cost-effectiveness estimates comparing software to CF were generally favourable to SF, except where the technologies were assumed to have the same diagnostic accuracy. The drivers of economic value of SF, comparative diagnostic accuracy and prevalence, are affected by unquantified uncertainty. Judgements on the economic value of SF require integration of the uncertainties over the clinical evidence with the overall cost-effectiveness.

Recommendations for further research

High-quality, sufficiently powered RCT evidence comparing SF biopsy with CF biopsy is required to address limitations from the existing evidence. Improved reporting of diagnostic accuracy outcomes would enable future syntheses to make use of a larger body of evidence.

Study registration

This trial is registered as PROSPERO CRD42022329259.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis programme (NIHR award ref: 135477) and is published in full in *Health Technology Assessment*; Vol. 28, No. 61. See the NIHR Funding and Awards website for further information.

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts 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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This article

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as award number NIHR135477. The protocol was agreed in May 2022. The draft manuscript began editorial review in December 2022 and was accepted for publication in June 2023. 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' manuscript 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 article.

This article presents independent research funded by the National Institute for Health and Care 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, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2024 Llewellyn *et al.* This work was produced by Llewellyn *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

