The clinical effectiveness and cost-effectiveness of the PROGENSA® prostate cancer antigen 3 assay and the Prostate Health Index in the diagnosis of prostate cancer: a systematic review and economic evaluation

Amanda Nicholson, James Mahon, Angela Boland, Sophie Beale,* Kerry Dwan, Nigel Fleeman, Juliet Hockenhull and Yenal Dundar

Liverpool Reviews and Implementation Group, University of Liverpool, Liverpool, UK

*Corresponding author

Declared competing interests of authors: none

Published October 2015
DOI: 10.3310/hta19870

Scientific summary

PROGENSA® PCA3 assay and the phi in the diagnosis of prostate cancer

Health Technology Assessment 2015; Vol. 19: No. 87
DOI: 10.3310/hta19870

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

Scientific summary

Background

Prostate cancer is a leading cause of mortality and morbidity. Approximately 40,000 new cases are diagnosed each year in the UK, and in 2011 10,793 deaths in the UK were attributed to the disease. The major risk factors for prostate cancer are increasing age, family history in a first-degree relative (brother or father) and race (higher risk in black men). The disease shows a strong inverse social gradient, being more common in more affluent social groups.

There is no single definitive test to identify men with prostate cancer. In cases in which prostate cancer could be the cause of presenting symptoms, the general practitioner carries out a number of tests. If, after carrying out this exploratory work, the general practitioner considers that there is a risk of prostate cancer, then the patient is referred to a hospital consultant to discuss the options for further tests.

The most commonly used test to detect prostate cancer is a transrectal ultrasonography-guided biopsy. However, this biopsy can miss cancers altogether, it may identify small, low-risk cancers that do not need to be treated but that may cause anxiety, it is uncomfortable (sometimes painful) and there can be complications for the patient (including blood in the urine, rectal bleeding and acute urinary retention). In some cases, when prostate cancer has not been confirmed by the initial biopsy, a second biopsy may be recommended. However, there is no guarantee that the second biopsy will find cancers missed by the first biopsy and further biopsies may still be required. Techniques such as multiparametric magnetic resonance imaging have been introduced into diagnostic practice. Such techniques improve the diagnostic performance of biopsies, as they help identify the location of prostate cancer abnormalities. However, multiparametric magnetic resonance imaging is not available in all hospitals.

The PROGENSA® prostate cancer antigen 3 (PCA3) assay (referred to as the PCA3 assay; Hologic Gen-Probe, Marlborough, MA, USA) and the Prostate Health Index (phi; Beckman Coulter Inc., Brea, CA, USA) are two new tests (a urine test and a blood test, respectively) that are designed to be used to help a clinician decide whether or not a second biopsy should be recommended. The purpose of this assessment was to evaluate the clinical effectiveness and cost-effectiveness of these tests, used in combination with existing tests, scans and clinical judgement, in the diagnosis of prostate cancer in men who are suspected of having malignant disease and in whom the results of an initial prostate biopsy were negative or equivocal. The perspective of the evaluation was the NHS in England and Wales.

Objectives

The key objectives of this assessment were to address the following questions:

1. How well do the PCA3 and [–2]pro-prostate-specific antigen (p2PSA) tests measure the substances they are intended to measure?
2. How might the addition of the PCA3 assay or phi contribute to the diagnosis of prostate cancer?
3. How might the addition of PCA3 assay or phi to current diagnostic strategies affect patient outcomes?
4. Would the addition of PCA3 assay or phi to current diagnostic strategies be cost-effective?

Methods

The research comprised two elements: an assessment of clinical effectiveness (addressing objectives 1, 2 and 3) and an assessment of cost-effectiveness (addressing objective 4). Literature searches to inform both elements were undertaken in May 2014.

Assessment of clinical effectiveness

Assessing the clinical effectiveness of the PCA3 assay and the phi in the diagnosis of prostate cancer involved three separate systematic reviews:

1. a review of the analytical validity (how well laboratory tests measure the substances they are intended to measure) of the intervention tests to assess how accurately the tests measure PCA3 score/p2PSA level present in a sample
2. a review of the clinical validity (accuracy of the diagnostic tests) of comparator and intervention pathways to assess how the addition of the PCA3 score or the phi might contribute to the diagnosis of prostate cancer
3. a review of the clinical utility of the intervention test pathways to evaluate how the addition of the intervention tests might affect patient outcomes, including long-term outcomes such as mortality and morbidity from prostate cancer and intermediate outcomes such as side effects from tests.

The methods used followed the systematic review principles outlined in the Centre for Reviews and Dissemination guidance for undertaking reviews in health care, the National Institute for Health and Care Excellence Diagnostic Assessment Programme manual and publications from the Cochrane diagnostic test accuracy methods working group. The review of analytical validity was informed by the principles outlined in the Agency for Healthcare Research and Quality methods guide and the Evaluation of Genomic Applications in Practice and Prevention initiative.

Assessment of cost-effectiveness

The cost-effectiveness assessment included two components: a systematic review of existing full economic evaluations and the development of a de novo health economic model.

The review of cost-effectiveness literature was conducted in line with the Centre for Reviews and Dissemination guidance for undertaking reviews in health care.

A de novo economic model was constructed using data from the clinical validity review. The External Assessment Group (EAG) model used values for derived specificities at defined sensitivity levels. By modelling defined sensitivities, the only difference between testing strategies was the number of biopsies required to identify a given number of cancers. The cost-effectiveness results were, therefore, driven by the differences in costs and quality-adjusted life-years (QALYs) losses that accrued in the different testing strategies (as a result of differences in numbers of biopsies performed). Incremental cost-effectiveness ratios were presented. Model input values for resource use, costs and utility values were extracted from published sources. The time horizon, in the base case, was 3 years (extended to 6 years in a scenario analysis), the model perspective was that of the UK NHS, and costs and benefits were discounted at a rate of 3.5%.

Results

Analytical validity review

To inform the assessment of the two assays, the EAG relied on data that have been published primarily by the manufacturers in the form of pack inserts and/or to support their submissions for regulatory approval. The review highlighted some important issues concerning the precision of PCA3 assay measurements and the requirements for storage and stability of samples for phi.

Clinical validity review

The key findings from the clinical validity review are as follows:

- Ten studies consider the comparison of clinical assessment versus clinical assessment + PCA3. The findings indicate that the implications of adding the PCA3 assay to clinical assessment are not clear and it is not possible to identify a single-threshold value for use in a clinical setting.
- Four studies consider the comparison of clinical assessment versus clinical assessment + phi. The findings indicate that the implications of adding phi to clinical assessment are not clear and it is not possible to identify threshold values for use in a clinical setting.
- Two studies consider the comparison of clinical assessment + magnetic resonance imaging (MRI) versus clinical assessment + MRI + PCA3. The findings indicate that the addition of the PCA3 assay to clinical assessment + MRI does not have a noticeable impact on discrimination.
- Only one study assesses the comparison of clinical assessment + MRI versus clinical assessment + MRI + phi. The findings indicate that the addition of phi to clinical assessment + MRI does not have a noticeable impact on discrimination.

Clinical utility review

The EAG did not identify any published papers that met the inclusion criteria for the clinical utility review.

Cost-effectiveness literature review

The EAG did not identify any published papers that met the inclusion criteria for the cost-effectiveness literature review.

Cost-effectiveness modelling

The key results from the base-case analyses are as follows:

- Clinical assessment versus clinical assessment + PCA3: clinical assessment dominates clinical assessment + PCA3 (i.e. clinical assessment costs less and generates more QALYs than clinical assessment + PCA3).
- Clinical assessment versus clinical assessment + phi: clinical assessment dominates clinical assessment + phi (i.e. clinical assessment costs less and generates more QALYs than clinical assessment + phi).
- Clinical assessment + MRI versus clinical assessment + MRI + PCA3: clinical assessment + MRI costs less but is less effective than clinical assessment + MRI + phi, and the incremental cost-effectiveness ratio per QALY gained for clinical assessment + MRI + phi is £5,418,366 compared with clinical assessment + MRI.
- Clinical assessment + MRI versus clinical assessment + MRI + phi: clinical assessment + MRI costs less but is less effective than clinical assessment + MRI + phi and the incremental cost-effectiveness ratio per QALY gained for clinical assessment + MRI + phi is £2,500,530 compared with clinical assessment + MRI.

Results from the sensitivity and scenario analyses show that, other than in one scenario which employed an unrealistic prostate-specific antigen (PSA) monitoring strategy, the incremental cost-effectiveness ratios that were generated to test model uncertainty are all above £20,000 per QALY gained. The probabilistic sensitivity analyses confirm that alternative testing strategies using any test in addition to clinical assessment are not cost-effective, although it should be noted that QALY loss associated with a biopsy was not varied in the probabilistic analyses.

Discussion

Strengths of the assessment

Although the assessment of analytical validity relied on data that have been published primarily by the manufacturers, the EAG considers that the analytical validity of the two tests has been comprehensively documented.

The clinical validity review includes results for a wide range of outcome measures for 10 different clinical comparisons. Its key strength is its focus on four clinically relevant comparisons, that is those studies reporting the addition of the PCA3 assay or phi to clinical assessment (with or without MRI).

The de novo economic model is based on the best available clinical validity evidence (identified through the systematic review) and captures the trade-off between the high upfront costs of diagnostic tests and the reduction in subsequent biopsies and their costs. It captures all of the main factors relevant to the decision problem and calculations are transparent.

Limitations of the assessment

The limitations of the clinical validity assessment are as follows:

- The review conclusions are over-reliant on findings from one study; of the 10 clinically relevant comparisons described in the 17 studies, data from one study are used in nine comparisons.
- The clinical relevance of many of the reported outcome measures is unclear.
- Many of the reported results for the clinical validity outcomes include neither standard errors nor confidence intervals.
- Descriptions of clinical assessment vary widely within reviewed studies.
- There was no consistent use, in the literature, of threshold values for either PCA3 score or phi.

The limitations of the economic assessment were as follows:

- There was a lack of generalisable clinical validity data to inform the economic model.
- The model was unable to capture and/or value all the key factors that might influence cost-effectiveness. The main area where information was lacking was in relation to utility decrements associated with prostate biopsies.

Uncertainties

Owing to the lack of published literature, the assessment was unable to address three clinical issues outlined in the final scope: detection of clinically insignificant cancer, optimal order of the tests and the effect of using different forms of reference standard (biopsy).

Further uncertainties, which relate to the economic model, include:

- the extent to which the model reflects NHS clinical practice
- the best way to model the most representative PSA monitoring strategy employed after a negative or equivocal biopsy in NHS clinical practice
- lack of clarity around the extent to which, in practice, clinicians prioritise sensitivity over specificity or vice versa.

Generalisability of the findings

The target population is not homogeneous but appears to comprise three subpopulations, namely those for whom a second biopsy is clearly indicated, those for whom a second biopsy is unnecessary and those for whom the need for a second biopsy is unclear. Most of the study populations described in the included studies comprise men who were referred for a second biopsy because of clinical suspicions and the criteria for referral varied between studies. The EAG considers, therefore, that it is not appropriate to apply the results of this review to all men with negative or equivocal biopsy results.

In addition, the representation of clinical assessment varied in the included studies. Although this may reflect clinical practice, in which clinical assessment is not standardised, it is difficult to meaningfully compare the results of studies which have markedly different representations of clinical assessment.

The reference standard (prostate biopsy) is an imperfect diagnostic tool as it does not detect all cancers. Without a gold standard that offers 100% specificity and 100% sensitivity, it is difficult to confidently assess the accuracy of competing diagnostic strategies.

Conclusions

Overall, the EAG considers that the analytical validity of the PCA3 assay and the phi has been comprehensively documented. The EAG identified some important issues relating to the precision of PCA3 assay measurements. Issues highlighted in relation to the use of the p2PSA assay were sample handling and the thermal stability of samples.

The clinical benefit of using the PCA3 assay and the phi in combination with existing tests, scans and clinical judgement in the diagnosis of prostate cancer in men who are suspected of having malignant disease and in whom the results of an initial prostate biopsy are negative or equivocal has not yet been confirmed. Furthermore, results from the cost-effectiveness analyses indicate that the use of these tests in the NHS for men who are suspected of having prostate cancer and have had a negative or equivocal initial biopsy would not be cost-effective.

Implications for service provision

A number of issues may affect the successful implementation of the assays in the NHS:

- PCA3 assay: the urine sample required for the PCA3 assay needs to be transferred to specialist transport tubes within 4 hours. Primary care staff may need some training for this requirement to be met. In addition, the published precision estimates for the PCA3 assay raise concerns about the interpretation and use of the PCA3 score for detecting prostate cancer.
- phi: blood samples for the p2PSA assay need to centrifuged and the serum separated within 3 hours. This time limit may pose challenges to implementing the test throughout the NHS. Furthermore, it is not clear whether or not blood samples taken in a primary care setting could be routinely transported to a laboratory and processed as required within 3 hours.

Suggested research priorities

Longitudinal end-to-end studies following men from initial investigation through to diagnosis and treatment of prostate cancer are required. Ideally, these studies would be randomised controlled trials with men allocated to different diagnostic test pathways after an initial negative or equivocal biopsy. However, descriptive data from observational cohorts following men over several years from initial referral onwards could address some unanswered issues.

Study registration

This study is registered as PROSPERO CRD42014009595.

Funding

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

Health Technology Assessment

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

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: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 14/03/01. The protocol was agreed in April 2014. The assessment report began editorial review in November 2014 and was accepted for publication in March 2015. 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.

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