PROTOCOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

The clinical and cost effectiveness of PROGENSA PCA3 Assay and the Prostate Health Index (phi) in the diagnosis of prostate cancer: a systematic review and economic evaluation

Completed 28th April 2014

DOES NOT CONTAIN CIC/AIC

UNIVERSITY OF LIVERPOOL
LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Table of contents

ABBREVIATIONS LIST

5ARI	5 alpha reductase inhibitors
ASAP	Atypical small acinar proliferation
AUC	Area under curve
DRE	Digital rectal examination
DW	Diffusion-weighted
EAG	External Assessment Group
fPSA	Free PSA
HGPIN	High grade prostatic intraepithelial neoplasia
MDT	Multidisciplinary team
MRI	Magnetic resonance imaging
mp MRI	Multiparametric MRI
mRNA	Messenger ribonucleic acid
PCA3	Prostate cancer gene 3
phi	Prostate Health Index
PSA	Prostate specific antigen
tPSA	Total PSA
RCT	Randomised controlled trial
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
TRUS	Transrectal ultrasound
WHO	World Health Organisation

1 TITLE OF PROJECT

Clinical and cost effectiveness of the PROGENSA PCA3 Assay and the Prostate Health Index in the diagnosis of prostate cancer: a systematic review and economic evaluation

2 NAME OF EXTERNAL ASSESSMENT GROUP (EAG) AND PROJECT LEAD

External Assessment Group:

Liverpool Reviews and Implementation Group (LR*i*G)

Project lead:

Dr Amanda Nicholson
Senior Research Fellow
University of Liverpool
Whelan Building
The Quadrangle
Brownlow Hill
Liverpool
L69 3GB
Tel: 0151 794 5447
Email: amanda.nicholson@liverpool.ac.uk

3 PLAIN LANGUAGE SUMMARY

Prostate cancer is the second most common cause of cancer-related death in men in the UK. It can be difficult to diagnose. Men with prostate cancer may present with clinical symptoms, or with a raised level of a protein called PSA. Increased PSA levels suggest a risk of prostate cancer but this test result is very non-specific and PSA levels can be raised in prostate conditions other than prostate cancer. Definitive cancer diagnosis usually requires a biopsy to obtain samples of prostate tissue which are then examined for cancerous cells. The prostate gland is situated deep within the pelvis area and it is difficult to access. Initial biopsy samples tend to be taken via the rectum, but this technique cannot access all areas of the prostate and there is a risk that the diseased area may be missed. Following negative or equivocal results from an initial biopsy, patients and doctors need to decide whether a second biopsy should be carried out. This decision requires balancing the discomfort and anxiety of a second biopsy with the risk of prostate cancer being undiagnosed and treatment delayed. Two new tests, the PROGENSA PCA3 Assay and the Prostate Health Index (phi), are available and results from these may help the decision making process. This study incudes a systematic review and an economic evaluation to assess the clinical and cost effectiveness (costs and benefits) of the use of the PROGENSA PCA3 Assay and phi, in combination with existing tests, scans and clinical judgement in the diagnosis of prostate cancer.

4 DECISION PROBLEM

4.1 Purpose of decision to be made

To evaluate the clinical and cost effectiveness of the PROGENSA PCA3 Assay and the Beckman Coulter Prostate Health Index (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.

4.2 Populations and relevant subgroups

4.2.1 Epidemiology

Prostate cancer is a leading cause of mortality and morbidity. Approximately 40,000 new cases are diagnosed each year in the UK[1] and in 2011 10,793 deaths in the UK were attributed to the disease.[2] The major risk factors are increasing age,[3] family history in a first degree relative (brother or father)[4] and race (higher in men of Afro-Caribbean origin).[5] The disease shows a strong inverse social gradient, being more common in more affluent social groups.[6] However, there is evidence that cancer is more likely to be at an advanced stage at diagnosis in the more deprived groups.[7]

The prognosis and natural history of prostate cancer vary depending on the extent of spread and the grade of cancer at diagnosis. The prognosis for men with disease localised to the prostate varies, and more aggressive changes on histopathology and higher prostate specific antigen (PSA) levels are associated with a worse prognosis.[8]

4.2.2 Current diagnostic practice

The recently updated National Institute for Health and Care Excellence (NICE) guideline[9] (Prostate cancer: diagnosis and treatment, CG175) summarises current best practice for the diagnosis and management of prostate cancer.

Men may initially present with clinical symptoms, such as difficulty with urination, or come to medical attention as the result of a raised PSA level. PSA is a protein produced in prostatic cells which can be elevated in men with prostate cancer. However, it is also raised in other benign prostatic conditions such as infections (prostatitis) and hypertrophy. A raised PSA is not, therefore, specific to the presence of cancer and not all men with prostate cancer have increased PSA levels. The decision whether or not to investigate for possible cancer is influenced by age as well as PSA level. Men in their 50s with PSA levels above 3 ng/ml are considered for further investigation, with cut-off levels being 4 ng/ml for men in their 60s and 5 ng/ml for men in their 70s.[10] The NICE guideline[9] recommends that the following factors should be taken into consideration when deciding to perform a biopsy: PSA level, digital rectal examination (DRE) findings, comorbidities, individual risk factors

such as increasing age, black Caribbean or black African ethnicity and family history. PSA level should not be used in isolation to guide a decision to biopsy.

Biopsy types

Diagnosis usually relies on obtaining a biopsy for histopathological examination of prostate tissue. The prostate gland is situated deep in the pelvis and it is not easy to visualise. Needle biopsies are obtained from the rectum under ultrasound control. The NICE guideline[9] recommends (p123) that prostate biopsies should be carried out following the procedure advocated by the Prostate Cancer Risk Management Programme (2006) *"Undertaking a transrectal ultrasound (TRUS) guided biopsy of the prostate"*.[11] This Programme advises that "the prostate should be sampled through the rectum unless there is a specific condition that prevents this" and also that "the scheme used at first biopsy should be a ten to twelve core pattern that samples the midlobe peripheral zone and the lateral peripheral zone of the prostate only" (section 11 Biopsy Scheme). These initial transrectal ultrasound (TRUS) biopsies are usually carried out under local anaesthetic as an out-patient or day case procedure.

TRUS biopsies are poor at accessing, and hence detecting, anterior, apical and central lesions.[12] Foci of cancerous cells may therefore be missed. This limits their usefulness and men who have prostate cancer may not have a positive result from an initial biopsy and one or more repeat biopsies may be taken before a diagnosis is confirmed.

The second biopsy may be another standard TRUS biopsy with eight to ten cores. However, more often, an increased number of samples are taken. Usual options are:

- saturation biopsy. A biopsy, which may be taken transrectally or transperineally, with an increased number of cores (minimum of 20). Men may prefer to have a general anaesthetic when undergoing a second biopsy, especially if they found the experience of their first biopsy to be uncomfortable and/or distressing
- template biopsy. 25 to 40 biopsy cores are taken transperineally using a template or grid to access more areas of the prostate, including anterior and apical zones. In the UK, this procedure is usually performed under general anaesthetic
- targeted biopsy. Information from a magnetic resonance imaging (MRI) scan is used to guide the biopsy to areas with disease (see MRI section 4.3.1 below).

Detection rates vary with the type of biopsy performed, number of cores taken and characteristics of the patient population; published estimates are 14-22% for first biopsy, 10-28% for second biopsy, and 5-10% for third biopsy.[13-16]

Prostate biopsies are painful and associated with side effects. Relatively common minor complications include haematospermia, haematuria and rectal bleeding which subsides after intervention, whilst

major complications, which are comparatively rare, include prostatitis, fever, urinary retention, epididymitis, and rectal bleeding for longer than 2 days.[17]

If cancer cells are detected, the histopathology report includes the Gleason score. The Gleason score (maximum score ten) describes the degree of abnormality of the tumour found in the biopsy and a score of six is indicative of low-grade cancer.[18] The proportion, or number, of samples with an abnormality detected is another important indicator of how widespread tumour cells are within the prostate.

Other abnormalities which may be reported on histopathology reports include:

- high grade prostatic intraepithelial neoplasia (HGPIN). This is a premalignant change in glands which has been shown to be associated with increased risk of invasive cancer elsewhere in the prostate
- atypical small acinar proliferation (ASAP). Atypical changes are present in cells but the pathologist is uncertain of their significance.

Clinically insignificant prostate cancer and risk stratification

The prognosis and natural history of prostate cancer vary with the extent of spread, and the grade of, cancer at diagnosis. Clinically insignificant prostate cancer can be defined as a cancer which will not affect the patient during the natural course of his lifetime, meaning that he is likely to die from other causes.[19] This means that active surveillance rather than other forms of treatment can be considered for some patients. The detection of these potentially clinically insignificant cancers on either initial or second biopsy is an important issue. It can lead to potentially invasive and unnecessary treatment as well as increased anxiety for men who live with a diagnosis of prostate cancer that may not affect their life expectancy.

The definitions of a clinically insignificant prostate cancer can be based on observed survival after radical prostatectomy. These pathology-based definitions require that the disease is restricted to the prostate, with a Gleason score of six, and some definitions include limits on the total tumour volume and/or largest individual tumour volume.[20,21] However, in clinical practice the challenge is to correctly identify men with clinically insignificant disease **before** any treatment or surgery. Various risk stratification systems for localised prostate cancer are in use,[22,23] and the current NICE guideline[9] follows the classification system based on the Harvard (D'Amico) system.[9,24] Men with localised prostate cancer are assessed as being at low, intermediate or high risk. Men with localised prostate cancer which is impalpable or restricted to one lobe only, a PSA level of less than 10 ng/ml and a Gleason score of six or below are considered to be at low risk of disease progression.[9] These men are offered active surveillance rather than other forms of treatment.

Recent data have suggested that these risk predictions may be inaccurate[23] and the NICE guideline[9] includes a research recommendation for further research into prognostic indicators.

4.3 Relevant comparators

This review is concerned with the decision to perform a second biopsy in men suspected of having prostate cancer who have a negative, or equivocal, initial biopsy.

The NICE guideline[9] reviewed the evidence of the efficacy of various prognostic factors used to determine the need for further investigation in men with a negative initial biopsy. The recommendations are as follows:

Recommendation 1: A core member of the urological cancer multi-disciplinary team (MDT) should review the risk factors of all men who have had a negative first prostate biopsy, and discuss with the man that the risk of prostate cancer is increased if any of the following risk factors is present:

- the biopsy shows HGPIN
- the biopsy shows atypical ASAP
- abnormal DRE.

Recommendation 2: To consider multiparametric MRI (mpMRI), using T2- and diffusion-weighted (DW) imaging, for men with a negative TRUS 10 to 12 core biopsy to determine whether another biopsy is needed.

Recommendation 3: Do not offer another biopsy if the mpMRI, using T2- and DW imaging, is negative, unless any of the risk factors listed in Recommendation 1 is present.

However, in clinical practice there may be considerable variation in the adherence to these recommendations.

4.3.1 Magnetic resonance imaging

MRI facilities and radiological expertise to allow guided biopsies are not available throughout the NHS. A survey of consultant radiologists from urological cancer MDTs is reported in the recent clinical guideline[9] and it shows that 73% of teams used MRI in the detection of prostate cancer. Of the teams using MRI, 50% use it prior to initial biopsy, 39% prior to second biopsy and 58% before subsequent biopsies (p 78). The exact role of MRI in guiding and informing subsequent biopsies varies. In cognitive targeting, knowledge of the MRI scan guides the freehand targeting of suspicious areas and requires no additional equipment. In direct MRI-guided biopsy, the biopsy is performed within an MRI tube. However, in fusion targeting, software is used to combine pre-acquired MRI-derived target with real-time TRUS imaging to guide the biopsy.[25,26]

In current NHS practice, MRI scans may not be performed for 6 to 12 weeks or longer after an initial biopsy, due to artefacts from bleeding. This has important time implications for the diagnostic pathways involving MRI after a negative or equivocal first biopsy and any subsequent treatment.

4.3.2 Other clinical factors

Clinicians and patients may use factors other than DRE and histopathology to inform the decision about whether or not a second biopsy should be carried out. These factors include:

- PSA level. The degree of elevation, especially in relation to estimated prostate volume, that can be expressed as PSA density[27]
- rising PSA levels (which can be expressed as PSA velocity (ng/ml increase over a time period))[28] or PSA doubling time
- patient's age
- family history.

Clinical decision-making involves a degree of subjective judgment to weigh up the information that is available. Most nomograms are statistically derived tools which may be used to describe the likely course of a disease using known variables such as diagnostic findings, age and treatment options. However, some nomograms (e.g. risk calculator number four from the Prostate Cancer Research Foundation,[29,30] the Prostate Cancer Prevention Trial (PCPT-CRC)[31] and Montreal nomograms[32]) can predict the result of a biopsy in men suspected of having prostate cancer. Some nomograms now include PCA3 levels.[33,34] It is not clear how often these tools are used to predict biopsy results in clinical practice but they may act as a proxy for clinical decision-making in the research setting.

4.4 Clear definition of the interventions

4.4.1 PCA3

The PROGENSA PCA3 Assay is an in-vitro nucleic acid amplification test that is intended for the quantitative determination of PCA3 messenger ribonucleic acid (mRNA) in urine. The PCA3 gene (previously known as DD3) is overexpressed in prostate cancer cells and is, therefore, a potential biomarker for tumour cells. Since 2002, methods to quantify the amount of PCA3 mRNA in urine have been available.[35-36] Prostatic cells are released into urine by prostatic massage but this leads to a general release of ribonucleic acid (RNA) and so the level of mRNA of another 'housekeeping' gene is needed to correct for the overall level of prostatic cells in the urine. The gene which encodes PSA (KK3 gene) has been selected as the 'housekeeping' gene as its mRNA expression is relatively constant in normal prostate cells, with only a weak down-regulation of PSA expression in prostate cancer cells. The PCA3 score report is a ratio of the PCA3 mRNA copies/mL to PSA mRNA copies/mL multiplied by 1000. The score can be used as a continuous measure but studies have used

cut-points of 20, 25 or 35 to identify men who are at higher risk of an underlying cancer.[37-40] The manufacturers of the PROGENSA PCA3 Assay have stated a threshold value (cut-point) of 25.

The assay analyses prostatic cells in the urine and needs to be preceded by DRE to apply pressure on each lobe of prostate to release prostatic cells and RNA. A first catch urine sample (of at least 2.5ml) is required after the examination and then 2.5ml of the sample is added to a transport tube containing a urine transport medium that triggers lysis of any prostatic cells and stabilises the RNA. The sample is then transported to a laboratory or can be kept frozen for 5 days before analysis.

The PROGENSA PCA3 Assay can be used with the Hologic Gen-Probe Direct Tube Sampling (DTS) 400, 800 and 1600 molecular laboratory systems. The PCA3 Assay is not compatible with other analyzers. Each PROGENSA PCA3 Assay kit is suitable for 2x100 reactions and includes reagents, controls and calibrators for both the PCA3 and PSA reactions. The PCA3 Assay package insert states that "PROGENSA PCA3 Assay should not be used for patients who are taking medications known to affect serum PSA levels such as finasteride (Proscar, Propecia), dutasteride (Avodart), and anti-androgen therapy (Lupron). The effect of these medications on PCA3 gene expression has not yet been evaluated".[41] Certain therapeutic and diagnostic procedures, including prostatectomy, radiation and prostate biopsy may affect the viability of prostatic tissue and, subsequently, an individual's PCA3 score. The effect of these procedures on assay performance has not yet been evaluated.

4.4.2 Prostate health index (phi)

The Prostate Health Index (phi) has been developed by Beckman Coulter to combine several different components of PSA with the aim of creating a sensitive index of risk of prostate cancer. Total PSA (tPSA) is measured in the blood stream where it occurs both unbound (free PSA, fPSA) and bound to other proteins (such as proteases). There is some evidence that the proportion of PSA that occurs unbound (%fPSA) is lower in men with cancer.[42,43] fPSA has been shown to include several isoforms including [-2]proPSA which is associated with cancerous cells. Beckman Coulter has developed an assay for [-2]proPSA (p2PSA) and the phi is calculated using the equation, ([–2]proPSA/free prostate specific antigen) × √ total PSA.[44,45]

The Beckman Coulter phi is designed for prostate cancer detection in men aged 50 years and older with tPSA levels between 2 ng/mL and 10 ng/mL and DRE findings that are not suspicious for cancer. The phi score is a continuous measure but it can be used in three or more categories to indicate the risk of prostate cancer being found on biopsy.[46] The manufacturer has suggested the following three categories of phi scores using WHO calibration: 0-20.9 (low risk); 21-39.9 (moderate risk); 40 and above (high risk). The manufacturer states that estimates of the risk of cancer being detected in biopsy

are 8.7% for men with a phi score in the low risk category, 20.6% for men in the moderate risk category and 43.8% for men in the high risk category.

The Beckman Coulter phi is not intended to be calculated using PSA or fPSA results from any other manufacturer's assay and the phi assay is only compatible with Beckman Coulter Access instruments (Access2, DxI600, DxI800, DxC600i, DxC680i, DxC800i, DxC880i). Other assays are available for tPSA and fPSA, and PSA values based on the World Health Organisation (WHO) standard and have been reported to be 20% lower than those from Hybritech standard.[47] A study comparing phi results using Hybritech calibration and WHO calibration for fPSA and tPSA has reported comparable phi results.[48]

The Beckman Coulter phi can be measured in blood samples but the [-2]proPSA molecule is not stable on coagulated blood. When left on a clotted sample at room temperature the [-2]proPSA concentration increases significantly after 3 hours, probably due to the degradation of other proPSA molecules. However, the analyte is stable in serum at room temperature. Therefore, it is important that the serum sample is prepared (separated from the clot by centrifugation) within 3 hours of taking a blood sample. Blood tests taken in a hospital with laboratory facilities on site will make this sample processing more feasible.

Information provided by the manufacturer states that the effect of medication for benign prostate hyperplasia (BPH), specifically the 5 alpha reductase inhibitors (5ARI), on the level of [-2]proPSA is not known. As a consequence, the phi results cannot be interpreted and should not be offered to patients receiving 5ARI medication.

4.5 Place of the intervention in the treatment pathway(s)

Based on the modelling possibilities outlined in the final scope (Appendix 1), the following diagnostic pathways for men suspected of having prostate cancer whose initial biopsy result was negative or equivocal will be considered:

Comparator pathways:

1. The use of established risk factors (histopathology results of initial biopsy, PSA level and DRE) to inform the decision to perform a second biopsy
2. The use of established risk factors (histopathology results of initial biopsy, PSA level and DRE) followed by mpMRI to inform the decision to perform a second biopsy.

Intervention pathways:

1. The use of PCA3/phi assays alongside established risk factors (histopathology results, PSA level and DRE) to inform the decision to perform a second biopsy

2a. The use of PCA3/phi assays alongside established risk factors (histopathology results, PSA level and DRE) to inform the decision to perform a mpMRI scan before second biopsy. If the mpMRI is positive a second biopsy would be performed

2b. The use of PCA3/phi assays alongside established risk factors (histopathology results, PSA level and DRE) to inform the decision to perform a second biopsy in men who have had a negative mpMRI scan.

These potential comparator and intervention diagnostic pathways are summarised in Figure 1. Men with prostate cancer detected on second biopsy will be offered treatment options depending on the grade and stage of the tumour. Men with a negative second biopsy will be assessed by their clinicians. There is however considerable uncertainty about the diagnostic pathways used in clinical practice and the availability of MRI scanning.

The sensitivity and specificity of the comparator and intervention pathways may differ with the type of second biopsy performed (extended, template, saturation or guided). Where possible, pathways will be modelled separately for these different biopsy types. However, if data from included studies in the systematic review are limited, data from various biopsy types may be pooled and differences explored in subsequent sensitivity analyses (see section 5.5.2).

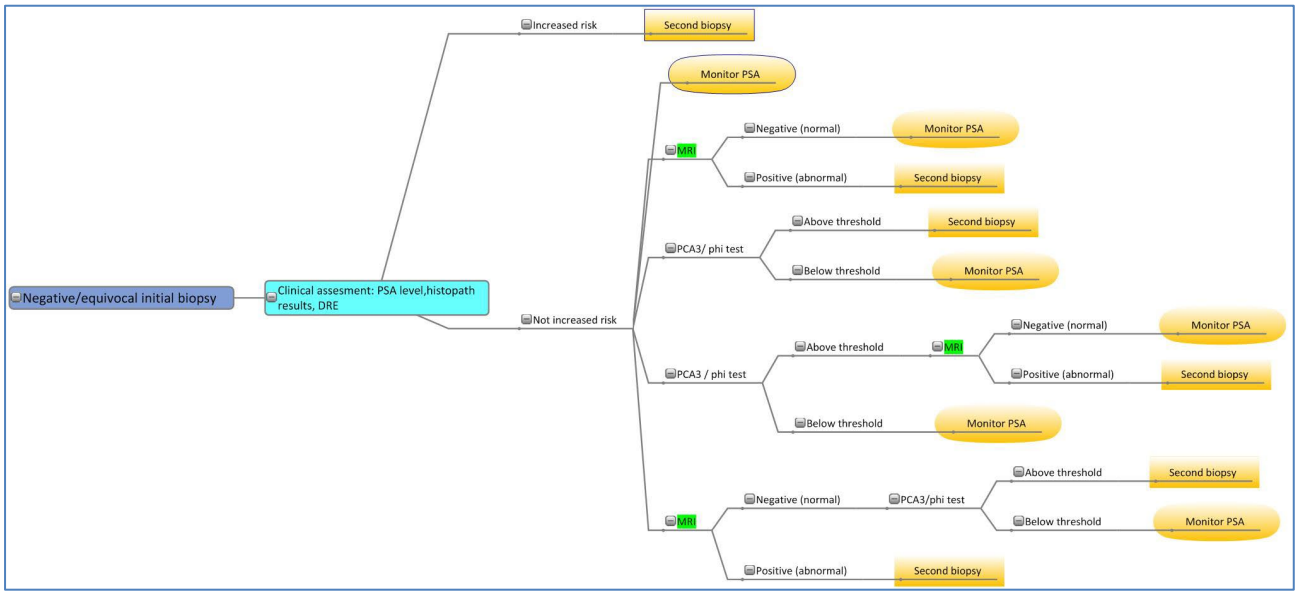

Figure 1. Diagnostic pathways for men after a negative or equivocal initial biopsy

4.6 Key factors to be addressed

4.6.1 Identification of clinically insignificant cancers

If the level of either PCA3[49] or phi[50] is shown to be associated with the histopathological grade (Gleason score) of tumours detected on subsequent biopsies, or with any other indicators of increased

risk such as spread, this may facilitate the targeting of biopsies away from lower risk cancers. This is an important potential contribution to the diagnostic process.

4.6.2 Clinical validity assessment

The intervention and comparator test pathways will be compared to the results of second biopsy which will form the reference standard for clinical validity studies. A diagnosis of prostate cancer on a biopsy will be considered as a positive reference standard result.

There are significant challenges to the interpretation of clinical validity studies included in this review:

- Both comparator and intervention pathways involve a sequence or combination of tests. Many clinical validity studies evaluate a single new test as a replacement for an existing test. However, in clinical practice new tests are often added to the existing battery of investigations. It will be important to attempt to establish precisely how the intervention tests are being used in a given study, i.e. whether they are used as add-on after existing tests or as a triage test before existing tests or alongside existing tests as part of a panel of investigations[51]
- The precise combination of tests used is not always clear in study reports. For example, targeted biopsies in effect incorporate a MRI scan with biopsy (the reference standard test)
- The PCA3 or phi threshold used to identify men at high risk may vary between studies. In other studies the results of intervention tests may be used as continuous variables. The optimum threshold will depend on the overall aim of the intervention. For example, whether the objective is to reduce the number of biopsies or to avoid missing cancers. The review will consider various threshold values and their impact on the accuracy and effectiveness of the interventions
- The reference standard (biopsy) does not detect all cancers and is considered to be imperfect. Furthermore, different types of biopsy may have different detection rates and the sensitivity and specificity of the intervention pathways may differ with type of second biopsy used.

4.7 Areas of agreement at the scoping workshop that are outside the scope of the appraisal and therefore do not require any detailed assessment

The External Assessment Group (EAG) will not consider the use of the intervention tests prior to initial biopsy or in the active surveillance of diagnosed prostate cancer.

There are no known differences in intervention test performance across different age groups or genotypes. Men receiving 5ARI medication, for example, finasteride (Proscar, Propecia), dutasteride (Avodart), or anti-androgen therapy (Lupron) will be excluded.

5 REPORT METHODS FOR ASSESSING THE OUTCOMES ARISING FROM THE USE OF THE INTERVENTIONS

This study of the clinical effectiveness of PCA3 and phi in the diagnosis of prostate cancer involves three separate systematic reviews:

- A review of the analytic validity of the intervention tests to assess how accurately the tests measure PCA3/phi level present in a sample
- A review of the clinical validity (diagnostic test accuracy) of comparator and intervention pathways to assess what the addition of the PCA3/phi assays contribute to the diagnosis of prostate cancer
- 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 in each review will follow the systematic review principles outlined in the Centre for Reviews and Dissemination (CRD) guidance for undertaking reviews in health care,[52] the NICE Diagnostic Assessment Programme manual[53] and publications from the Cochrane Collaboration diagnostic test accuracy methods working group.[54] The review of analytic validity will be informed by the principles outlined in the Agency for Healthcare Research and Quality (AHRQ) methods guide[55] and the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative.[56]

5.1 Search strategy

5.1.1 Electronic databases

The following databases will be searched for eligible studies:

- CDSR (Cochrane Database of Systematic Reviews)
- DARE (Database of Abstracts of Reviews of Effectiveness)
- MEDION database for related diagnostic test accuracy reviews
- ARIF (Aggressive Research Intelligence Facility) database
- MEDLINE
- EMBASE
- CENTRAL (Cochrane Central Register of Controlled Trials Health Technology Assessment (HTA) database)
- ISI Web of Science

No study design filters will be applied and non-English language reports will be excluded. All databases will be searched from 2000 until the latest available version.

Trial and research registers will be searched for ongoing trials and reviews including:

- Clinicaltrials.gov
- *meta*Register of Controlled Trials and ISRCTN Register
- WHO International Clinical Trials Registry Platform
- Prospero systematic review register.

Details of the draft search for Medline can be found in Appendix 2.

5.1.2 Searching other resources

Backward and forward citation searching will be undertaken on key review articles and on eligible studies identified from the initial searches. Investigators known to be involved in previous studies will be contacted to enquire about ongoing or unpublished studies. The online resources of various health services research agencies and professional societies will be consulted via the internet.

5.2 Study selection

The citations identified will be assessed for inclusion through two stages. Firstly, two reviewers will independently scan all the titles and abstracts identified by the searching exercise to identify the potentially relevant articles to be retrieved. Full text copies of the selected studies will subsequently be obtained and assessed independently by two reviewers for inclusion using the inclusion criteria outlined below. Any disagreements will be resolved by discussion at each stage, and, if necessary, a third reviewer will be consulted.

5.2.1 Studies for review of analytic validity

This review will focus on studies that address the ability of the intervention test to accurately and reliably measure the target analyte. Inclusion criteria are presented in Table 1.

Table 1 Inclusion criteria (analytic validity)

Patient population	All adult men
Outcomes	• Measures of consistency and accuracy between, and within, laboratories such as coefficient of variation • Sensitivity and specificity against external standard • Assay robustness • Test failure rate
Study design	All study designs including collaborative studies, external proficiency testing, peer-reviewed repeatability studies, internal reports and manufacturer data

5.2.2 Studies for review of clinical validity

Within-study (direct) comparisons

The preferred data for this review are from within-study (direct) comparisons of intervention and comparator test pathways. Due to the uncertainty about the diagnostic pathways used in NHS clinical practice and the availability of MRI scanning facilities, the EAG will include all studies with a direct comparison of PCA3 and/or phi with any **one or more** of following component comparator tests:

- individual clinical risk factors such as age, DRE
- standard clinical judgment/nomograms
- PSA levels
- MRI results: T2-MRI/DW-MRI.

Given the likely variation and uncertainty about the combination of tests analysed and reported in relevant studies, eligible studies may have used the intervention tests (PCA3 Assay or phi) as replacement, add-on or triage tests to the comparator tests. Studies that have directly compared the performance of the PCA3 Assay with that of phi, with or without other comparators, will also be included.

Possible study designs for within-study comparisons include paired designs (where intervention test, comparator tests and reference standard are all performed on the same group of participants) and unpaired designs (trials in which participants are randomised to receive either the intervention or comparator test and then all participants undergo the reference standard test).

Between-study (indirect) comparisons

The EAG will consider carrying out analyses based on between-study (indirect) comparisons of the intervention tests with comparator tests. In these analyses, estimates of the accuracy of the tests will be derived from studies in which the intervention or comparator alone was compared to reference standard. Estimates of the clinical validity of the intervention or comparator tests from good quality systematic reviews and meta-analyses will be sought. The EAG acknowledges that results from indirect analyses are less reliable than those using direct comparison studies as any differences in clinical validity could be due to confounding factors between studies. However, these indirect studies will also provide data that will confirm (or refute) the sensitivity and specificity estimates for the comparator pathways that were obtained from the smaller direct comparison studies.

Inclusion criteria are presented in Table 2.

Table 2 Inclusion criteria (clinical validity)

Patient population	Men suspected of having prostate cancer who have had one negative or equivocal biopsy. The review is restricted to studies where at least six cores were taken in initial biopsy
Intervention	Diagnostic test or test pathway including PCA3 and/or phi
Comparator	Diagnostic test or test pathway without PCA3 or phi and including one or more of following comparator tests • individual clinical risk factors such as age, DRE • standard clinical care/nomograms • PSA levels • MRI results: T2-MRI/ DW-MRI Studies that have directly compared the performance of PCA3 with that of phi, with or without other comparators, will also be included
Reference standard	Eligible studies must compare the performance of comparator or intervention pathways to a histological analysis of prostatic tissue. This may be obtained from a second prostatic biopsy or from prostatectomy specimen Biopsy must have taken place within 1 year of the intervention test All types of second biopsy will be included: • repeat standard TRUS biopsy • saturation • template • MRI targeted biopsies • use of prostatectomy specimens
Outcomes	Clinical validity data may be presented in various ways in eligible studies. Studies that report any of the following will be included: • estimates of the intervention or comparator test (Means and standard deviation (SD), proportion positive) in men with positive and negative results on second biopsy • specificity and sensitivity for different cut-off points of PCA3, phi or PSA • comparison of AUC for different tests or test combinations • gain in sensitivity and specificity estimates by adding intervention test as derived from ROC curves • results of logistic regression analyses These studies may also provide outcome data: • test failure rate • adverse effects of test or subsequent biopsies • risk group and stage of cancers detected
Study design	<u>Studies reporting within-study comparison of interventions/comparators:</u> • paired design. Cross-sectional or longitudinal studies in which intervention test(s), comparator test(s) and reference standard test were performed in the same group of people • unpaired design. Trials in which people were randomised to either the intervention or comparator test(s) and then all received the reference standard test <u>Studies for inclusion in between-study comparisons of interventions/ comparators:</u> • systematic reviews and meta-analyses of the clinical validity of the intervention or any of the comparator tests

5.2.3 Studies for review of clinical utility

This section of the assessment aims to evaluate how the addition of the intervention tests affects patient outcomes, including long-term outcomes such as mortality and morbidity from prostate

cancer, and intermediate outcomes, such as side effects from tests. “End-to-end” or “test-to-treatment” studies examining the effect of the use of tests on subsequent patient care and outcomes would give the highest level evidence but such studies are unlikely to be available. The ideal design for these studies would be randomised controlled trials (RCTs) in which people were randomised to the index and comparator test pathways and followed through subsequent treatment. If no randomised trials are available, the EAG will use data from any available design including observational cohorts and/or patient surveys as these non-randomised studies may also provide data on outcomes for men who have followed different diagnostic pathways. Priority will be given to higher level evidence such as RCTs. Inclusion criteria are presented in Table 3.

Studies of clinical validity, as described in section 5.2.2, may report some clinical outcomes such as adverse effects of biopsy or patient reported anxiety or distress.

Table 3 Inclusion criteria (clinical utility)

Patient population	Men suspected of having prostate cancer who have had one negative or equivocal biopsy
Intervention	PCA3/phi or test pathway including PCA3 or phi
Comparator	Diagnostic test or test pathway without PCA3/phi and including one or more of following comparator tests: • individual clinical risk factors such as age, DRE • standard clinical care/nomograms • PSA levels • MRI results: T2-MRI/DW-MRI
Outcomes	Intermediate outcomes (diagnostic process outcomes): • diagnostic test accuracy • test failure rate • time to true positive diagnosis • number of repeat biopsies required • grade and stage of cancers detected Clinical outcomes: • morbidity and mortality from biopsies • morbidity and mortality from treatment of diagnosed cancer • adverse events from false test results including from treatment of clinically insignificant prostate cancer • health related quality of life Patient reported outcomes: • patient anxiety associated with undergoing a biopsy (initial and repeated biopsies), waiting for diagnosis and living with the diagnosis of a clinically insignificant prostate cancer • patient distress and sequelae associated with the detection of clinically insignificant prostate cancer
Study design	All study designs eligible including: • randomised controlled trials • observational cohorts • patient surveys • studies of clinical validity

5.3 Data extraction strategy

Data will be extracted from eligible studies independently by two reviewers using a paper-based data extraction form designed for each type of review. Draft forms are included in Appendix 3. These forms will be reviewed after data have been extracted from the first three studies included in each review. If there are duplicate eligible publications from the same study a composite dataset will be created from all the relevant publications. If time permits, authors (and sponsors) of the studies will be contacted for missing data.

Analytic validity review: Data extraction will include details of source population, numbers of samples, specific methods/platforms evaluated, number of positive samples and negative controls tested, as well as reported results.

Clinical validity review: In studies reporting a combination of tests, particular attention will be paid to:

- how the intervention and comparator tests have been used: replacement, add-on, triage or not stated in the paper
- type of second biopsy (reference standard) used
- definition of positive biopsy, including grade and stage of tumour detected
- threshold values used.

The available data on all clinical validity outcomes reported will be recorded including:

- 2x2 tables of true positive (TP), false positive (FP), false negative (FN) and true negative (TN)
- sensitivity, specificity, positive and negative predictive value, positive and negative likelihood ratios
- area under curve (AUC) and sensitivity and specificity derived from receiver operating characteristic (ROC) curves
- odds ratios for logistic regression.

If multiple thresholds are used and reported, outcomes will be recorded for each threshold.

If different combinations or sequence of tests are reported, outcomes will be recorded for each combination or sequence.

If accuracy data are presented separately for different risk groups (grades or stages) of tumour detected, outcomes will be recorded for different risk groups.

Clinical utility review: Data extracted from studies of patient outcomes will include details of patient population, study design, intervention and control groups. The available data on all clinical outcomes will be reported including:

- morbidity and mortality from biopsies
- morbidity and mortality from treatment of diagnosed cancer
- health related quality of life
- patient reported outcomes.

5.4 Quality assessment strategy

Quality assessment tools appropriate to the study design will be employed. Quality assessment will be undertaken independently by two reviewers and any disagreements resolved by a third reviewer.

5.4.1 Studies of analytic validity:

Quality assessment will be informed by the checklist proposed by Tuetsch[56] and will include:

- quality of description of test undertaken
- range of samples / study population tested representative of routine use
- definition of correct answer
- reporting of test failures.

5.4.2 Studies of clinical validity:

The QUADAS-2,[57] a modified version of the Quality Assessment of Diagnostic Accuracy Studies tool,[58,59] will be used to assess the quality of included studies. This tool considers four domains: patient selection, index tests(s), reference standard, flow and timing. The tool content has been tailored to meet the requirements for this review and a copy of the tool may be found in Appendix 4. The following issues are of particular importance to this review:

- differences in reference standard between test arms
- incorporation bias of intervention results into reference standard (for instance, MRI results can be used to inform biopsy sampling)
- spectrum bias, i.e. patients selected on the basis of their PSA levels.

5.4.3 Clinical utility studies

The quality of studies reporting clinical outcomes will be assessed using quality assessment tools appropriate to the study design. Potential instruments include the Cochrane risk of bias tool[52] for RCTs and the Downs and Black checklist[60,61] for other study designs.

5.5 Methods of analysis/synthesis

5.5.1 Analytic validity

A narrative summary of the estimates available will be presented; a meta-analysis of these results is not planned.

5.5.2 Clinical validity

The extent of data synthesis undertaken will depend on the data available within the primary studies.

Meta-analysis techniques are best developed for estimates of sensitivity and specificity[54] and in this review would require measures of sensitivity and specificity when using PCA3/phi at particular thresholds. The diagnostic accuracy of a combination and/or sequence of tests is a key feature of this review. Ideally these data would be reported within a primary study and be presented in an appropriate format.[51,62] For instance, cross-tabulation of the results of all of the tests included in a sequence would be presented separately for men who were positive and negative on second biopsy. If at least three comparable estimates of sensitivity and specificity for **equivalent intervention and comparators test pathways are identified, which use the same threshold value and the same reference standard,** meta-analysis will be undertaken as recommended in the NICE Diagnostic Assessment Programme manual and Cochrane diagnostic test accuracy handbook.[53,54] The use of bivariate models will be considered by the EAG as these are suitable for assessing summary sensitivity and specificity at a common threshold, or at several different common thresholds.[54] Revman 5.2 and Stata 12 software (*metandi* or *xtmelogit* command)[63] may be used.

However, such detailed sensitivity and specificity data for comparable pathways may be sparse. More often, combinations of tests are presented in a diagnostic risk prediction model[64] using logistic regression. Increments in AUC and improvements in sensitivity at a given specificity, or vice versa, are often presented. PCA3 and phi values may be entered as continuous variables, which means that it is not possible to relate any improvements in accuracy to the use of a particular threshold value. Methods for meta-analysis of ROC from such prediction models are less developed than for sensitivity and specificity and narrative synthesis may be required.

Sensitivity analyses

If data are available, sensitivity analyses will be performed to assess the impact of the following variables on the accuracy of tests:

- type of second biopsy (saturation, template or guided)
- threshold value used or results used as a continuous variable.

The clinical validity results for each test pathway will be stratified by biopsy type and threshold value and, if possible, heterogeneity will be investigated by adding a covariate to bivariate or other summary model.

If data are available for the different risk groups (grades or stages) of tumour detected at second biopsy, the accuracy of the test pathways will be summarised for these different risk groups.

5.5.3 Clinical utility

The decision to pool clinical utility study results will depend on the extent of the methodological and clinical heterogeneity identified in the included studies. Results from single arm studies (observational and/or non-comparative studies) may be considered for pooling should they be relevant to the research objectives. Outcome data may include time-to-event (survival), binary outcomes, continuous measures (such as anxiety or distress) and choice of meta-analysis. Statistical heterogeneity will be assessed using chi-square and I^2 statistics. Depending on the level of clinical and statistical heterogeneity, subgroup and sensitivity analyses will be explored.

If no end-to-end studies are identified, a linked evidence approach will be used to combine data on clinical outcomes from other studies with the summary clinical validity data. These data will be derived from a range of study designs including systematic review, RCTs of treatment options, comparative cohort studies and descriptive cohort studies.

6 REPORT METHODS FOR SYNTHESIZING EVIDENCE OF COST EFFECTIVENESS

6.1 Identifying and systematically reviewing published cost effectiveness studies

The search strategy detailed in section 5 will be used to identify studies examining the cost effectiveness of using PROGENSA PCA3 Assay and phi as aids to diagnosis in people undergoing investigations for suspected prostate cancer. Other searching activities, including electronic searching of online health economic journals and contacting experts in the field will also be undertaken. Full details of the search process will be presented in the final report.

Titles and abstracts will be examined for inclusion by two reviewers independently. Potentially relevant studies will then be obtained in full text and examined more carefully by two independent reviewers using pre-specified inclusion/exclusion criteria, details of which will be described in the final report. Any disagreement will be resolved by consensus, and if necessary, a third reviewer will be consulted.

Only full economic evaluations (assessing both outcomes and benefits) will be included in the review. However, to supplement findings, additional information on costs and benefits will be collated and discussed narratively as appropriate.

Data from the full economic evaluations meeting the inclusion criteria will be extracted into structured tables and will include, but not be limited to, the variables set out in Appendix 3. The

quality of the included studies will be assessed using the critical appraisal checklist for economic evaluations proposed by Drummond and colleagues.[65]

6.2 Development of a health economic model

6.2.1 Model pathways

The NICE scope, provided in Appendix 1, details three diagnostic pathway options:

- PCA3 or phi testing is carried out after an initial negative biopsy to determine, in combination with information on other risk factors, if a second biopsy is needed followed by standard care as described in CG175[9]
- PCA3 or phi testing is carried out after an initial negative biopsy to determine, in combination with information on other risk factors, if an mpMRI is needed. If the mpMRI is positive, a biopsy would be performed followed by standard care as described in CG175[9]
- PCA3 or phi testing is carried out after both an initial negative biopsy and a negative mpMRI to determine, in combination with information on other risk factors, if a second biopsy is needed followed by standard care as described in CG175.[9]

There appears to be considerable variability in the diagnostic pathways used in clinical practice and, in particular, the degree to which MRI scanning is an available option in the UK NHS. Furthermore, it is anticipated that the effectiveness evidence for the PCA3 Assay and phi may be limited and that the data that are available may be difficult to include in an economic model because of the way clinical effectiveness is reported.

Modelling will be restricted to those scenarios that are particularly relevant to current NHS clinical practice and for which credible evidence is available. To identify such scenarios the EAG will, once an initial assessment of the effectiveness evidence has been carried out, circulate a brief paper detailing the diagnostic pathways for which credible evidence exists. This paper will highlight where data are available, how robust those data are, and where there are data gaps. It will be circulated, via email, to members of the Assessment Sub-Group (ASG), NICE personnel and the EAG's clinical advisors. Recipients will be asked for their opinion on which pathways should be modelled. A final decision about which pathways to include in the economic model will be made based on the views received from the experts and following discussions with colleagues at NICE.

6.2.2 Model structure

Necessary choices and definitions regarding the structure of the model will be influenced by findings from the literature as well as from expert clinical advice. Targeted literature searches will be carried out to identify peer reviewed studies published since 2009 which report on cost-effectiveness models considering different interventions used in the diagnostic pathways for prostate cancer. It is

anticipated that, in particular, findings from the literature will inform the development of the post diagnosis element of the economic model.

The patient population considered in the model will be men suspected of having prostate cancer who have had one negative or equivocal biopsy. The economic appraisal will be undertaken from the perspective of the NHS and Personal Social Services. Where possible, published national costs will be used, for example NHS Reference Costs[66] and Unit Costs for Health and Social Care.[67] Where national costs are not available, costs will be based on estimates from the published literature. If there are no data in the published literature, values will be sought from experts. The model time horizon will be set to patient life-time (estimated to be 30 years in the base case) and both costs and benefits will be discounted at 3.5%.

Model results will be presented as incremental cost per quality-adjusted life year (QALY) ratios. The literature will be searched for appropriate utility values to use in the model. The search will include citation searches of key papers that have been quoted in recent economic models. To illustrate, Mowatt et al[71] and the recent NICE clinical guideline[9] have used data from Korfage et al,[68] Krahn et al,[69] Shimizu et al[70] and Volk et al.[71]

Appropriate sensitivity analyses will be undertaken to assess the robustness of the model results to realistic variations in the levels of the underlying data. Where the overall results are sensitive to a particular variable, the sensitivity analysis will analyse the exact nature of the impact of variations.

Imprecision in the principal model cost-effectiveness results with respect to key parameter values will be assessed by use of techniques compatible with the modelling methodology deemed appropriate to the research question (e.g. multi-way sensitivity analysis, cost effectiveness acceptability curves).

7 OTHER INFORMATION

7.1 Handling information from the companies

All data submitted by the manufacturers/sponsors will only be considered if received by the EAG before 01/09/2014. Data arriving after this date will not be considered. Any data that meet the inclusion criteria stated will be extracted and quality assessed as stated in the methods section of this protocol.

Any 'commercial in confidence' data provided by manufacturers, and specified as such, will be highlighted in blue and underlined in the assessment report (followed by company name in parentheses). Any 'academic in confidence' data provided by manufacturers, and specified as such, will be highlighted in yellow and underlined in the assessment report. All confidential data used in the cost effectiveness models will also be highlighted.

7.2 Competing interests of authors

None of the authors have any competing interests

7.3 Project timetable

Table 4 Project timelines

Activity	Date to be completed
Draft protocol submission	31st March 2014
Final protocol submitted	25th April 2014
Beginning of review process	1st May 2014
Literature search and assessment of papers for review	May 2014
Data extraction	June 2014
Data synthesis and economic modelling	July, August, September 2014
Draft report for review	Late September 2014
Report submitted	17th October 2014

8 REFERENCES

1. Cancer Research UK. Prostate cancer incidence statistics. 2013. Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/prostate/incidence/. [Accessed March 2014].
2. Cancer Research UK. Prostate cancer mortality statistics. 2013. Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/prostate/mortality/uk-prostate-cancer-mortality-statistics. [Accessed March 2014].
3. Patel AR, Klein EA. Risk factors for prostate cancer. Nat Clin Pract Urol. 2009; 6:87-95.
4. Goh CL, Schumacher FR, Easton D, Muir K, Henderson B, Kote-Jarai Z*, et al.* Genetic variants associated with predisposition to prostate cancer and potential clinical implications. J Intern Med. 2012; 271:353-65.
5. Ben-Shlomo Y, Evans S, Ibrahim F, Patel B, Anson K, Chinegwundoh F*, et al.* The risk of prostate cancer amongst black men in the United Kingdom: the PROCESS cohort study. Eur Urol. 2008; 53:99-105.
6. National Cancer Intelligence Network. Cancer incidence and survival by major ethnic group, England, 2002-2006. 2009. Available from: http://publications.cancerresearchuk.org/downloads/product/CS_REPORT_INCSURV_ETHNIC.pdf. [Accessed March 2014].
7. Lyratzopoulos G, Abel GA, Brown CH, Rous BA, Vernon SA, Roland M*, et al.* Socio-demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer. Ann Oncol. 2013; 24:843-50.
8. American Cancer Society. How is prostate cancer staged? 2013. Available from: http://www.cancer.org/cancer/prostatecancer/detailedguide/prostate-cancer-staging. [Accessed March 2014].
9. National Collaborating Centre for Cancer. Prostate cancer: diagnosis and treatment. 2014. Available from: http://guidance.nice.org.uk/CG175.
10. Prostate Cancer Risk Management Programme. Information for primary care; PSA testing in asymptomatic men. 2010. Available from: http://www.cancerscreening.nhs.uk/prostate/pcrmp02.pdf. [Accessed March 2014].
11. Prostate Cancer Risk Management Programme. Undertaking a transrectal ultrasound guided biopsy of the prostate. 2006. Available from: http://www.cancerscreening.nhs.uk/prostate/pcrmp01.pdf. [Accessed March 2014].
12. Norberg M, Egevad L, Holmberg L, Sparen P, Norlen BJ, Busch C. The sextant protocol for ultrasound-guided core biopsies of the prostate underestimates the presence of cancer. Urology. 1997; 50:562-6.
13. Djavan B, Waldert M, Zlotta A, Dobronski P, Seitz C, Remzi M*, et al.* Safety and morbidity of first and repeat transrectal ultrasound guided prostate needle biopsies: results of a prospective European prostate cancer detection study. J Urol. 2001; 166:856-60.
14. Lujan M, Paez A, Santonja C, Pascual T, Fernandez I, Berenguer A. Prostate cancer detection and tumor characteristics in men with multiple biopsy sessions. Prostate Cancer Prostatic Dis. 2004; 7:238-42.
15. Mian BM, Naya Y, Okihara K, Vakar-Lopez F, Troncoso P, Babaian RJ. Predictors of cancer in repeat extended multisite prostate biopsy in men with previous negative extended multisite biopsy. Urology. 2002; 60:836-40.

16. Scattoni V, Raber M, Capitanio U, Abdollah F, Roscigno M, Angiolilli D*, et al.* The optimal rebiopsy prostatic scheme depends on patient clinical characteristics: results of a recursive partitioning analysis based on a 24-core systematic scheme. Eur Urol. 2011; 60:834-41.
17. Eichler K, Wilby J, Hempel S, Myers L, Kleijnen J. Diagnostic value of systematic prostate biopsy methods in the investigation for prostate cancer. A systematic review. 2005. Available from: https://www.york.ac.uk/inst/crd/CRD_Reports/crdreport29.pdf. [Accessed March 2014].
18. University of Liverpool. Towards a Consensus Protocol on Prostate Biopsies: Indications, Techniques and Assessment. 2003. Available from: http://www.cancerscreening.nhs.uk/prostate/conference-report.pdf. [Accessed April 2014].
19. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. JAMA. 2005; 293:2095-101.
20. Stamey TA, Freiha FS, McNeal JE, Redwine EA, Whittemore AS, Schmid HP. Localized prostate cancer. Relationship of tumor volume to clinical significance for treatment of prostate cancer. Cancer. 1993; 71:933-8.
21. Wolters T, Roobol MJ, van Leeuwen PJ, van den Bergh RCN, Hoedemaeker RF, van Leenders GJLH*, et al.* A Critical Analysis of the Tumor Volume Threshold for Clinically Insignificant Prostate Cancer Using a Data Set of a Randomized Screening Trial. J Urol. 2011; 185:121-5.
22. Rodrigues G, Warde P, Pickles T, Crook J, Brundage M, Souhami L*, et al.* Pre-treatment risk stratification of prostate cancer patients: A critical review. Can Urol Assoc J. 2012; 6:121-7.
23. Shaw GL, Thomas BC, Dawson SN, Srivastava G, Vowler SL, Gnanapragasam VJ*, et al.* Identification of pathologically insignificant prostate cancer is not accurate in unscreened men. Br J Cancer. 2014.
24. D'Amico AV, Whittington R, Malkowicz S, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. 1998; 280:969-74.
25. Le JD, Huang J, Marks LS. Targeted prostate biopsy: value of multiparametric magnetic resonance imaging in detection of localized cancer. Asian J Androl. 2014: Available from: http://www.asiaandro.com/aoparts.asp?doi=10.4103/1008-682X.122864.
26. Marks L, Young S, Natarajan S. MRI–ultrasound fusion for guidance of targeted prostate biopsy. Current Opinion in Urology. 2013; 23:43-50
27. Benson MC, Whang IS, Pantuck A, Ring K, Kaplan SA, Olsson CA*, et al.* Prostate specific antigen density: a means of distinguishing benign prostatic hypertrophy and prostate cancer. J Urol. 1992; 147:815-6.
28. Thanigasalam R, Mancuso P, Tsao K, Rashid P. Prostate-specific antigen velocity (PSAV): a practical role for PSA? ANZ J Surg. 2009; 79:703-6.
29. The Prostate Cencer Research Foundation R. Risk calculators. Available from: http://www.prostatecancer-riskcalculator.com/seven-prostate-cancer-risk-calculators. [Accessed April 2014].
30. van Vugt HA, Roobol MJ, Kranse R, Maattanen L, Finne P, Hugosson J*, et al.* Prediction of prostate cancer in unscreened men: external validation of a risk calculator. Eur J Cancer. 2011; 47:903-9.

31. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, *et al.* Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst. 2006; 98:529-34.
32. Karakiewicz PI, Benayoun S, Kattan MW, Perrotte P, Valiquette L, Scardino PT, *et al.* Development and validation of a nomogram predicting the outcome of prostate biopsy based on patient age, digital rectal examination and serum prostate specific antigen. J Urol. 2005; 173:1930-4.
33. Auprich M, Haese A, Walz J, Pummer K, de la Taille A, Graefen M, *et al.* External validation of urinary PCA3-based nomograms to individually predict prostate biopsy outcome. Eur Urol. 2010; 58:727-32.
34. Perdona S, Cavadas V, Di Lorenzo G, Damiano R, Chiappetta G, Del Prete P, *et al.* Prostate cancer detection in the "grey area" of prostate-specific antigen below 10 ng/ml: head-to-head comparison of the updated PCPT calculator and Chun's nomogram, two risk estimators incorporating prostate cancer antigen 3. Eur Urol. 2011; 59:81-7.
35. Durand X, Moutereau S, Xylinas E, de la Taille A. Progensa PCA3 test for prostate cancer. Expert Rev Mol Diagn. 2011; 11:137-44.
36. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeney LA, Aalders TW, *et al.* DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. Cancer Res. 2002; 62:2695-8.
37. Gittelman MC, Hertzman B, Bailen J, Williams T, Koziol I, Henderson RJ, *et al.* PCA3 molecular urine test as a predictor of repeat prostate biopsy outcome in men with previous negative biopsies: a prospective multicenter clinical study. J Urol. 2013; 190:64-9.
38. Haese A, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PF, *et al.* Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. Eur Urol. 2008; 54:1081-8.
39. Pepe P, Fraggetta F, Galia A, Skonieczny G, Aragona F. PCA3 score and prostate cancer diagnosis at repeated saturation biopsy. Which cut-off: 20 or 35? Int Braz J Urol. 2012; 38:489-95.
40. Sokoll LJ, Ellis W, Lange P, Noteboom J, Elliott DJ, Deras IL, *et al.* A multicenter evaluation of the PCA3 molecular urine test: pre-analytical effects, analytical performance, and diagnostic accuracy. Clin Chim Acta. 2008; 389:1-6.
41. Gen-probe. Progensa PCA3 asssay package insert. Available from: http://gen-probe.com/pdfs/pi/501377RevB.pdf. [Accessed April 2014].
42. Lee R, Localio AR, Armstrong K, Malkowicz SB, Schwartz JS. A meta-analysis of the performance characteristics of the free prostate-specific antigen test. Urology. 2006; 67:762-8.
43. Roddam AW, Duffy MJ, Hamdy FC, Ward AM, Patnick J, Price CP, *et al.* Use of prostate-specific antigen (PSA) isoforms for the detection of prostate cancer in men with a PSA level of 2-10 ng/ml: systematic review and meta-analysis. Eur Urol. 2005; 48:386-99; discussion 98-9.
44. Jansen FH, van Schaik RH, Kurstjens J, Horninger W, Klocker H, Bektic J, *et al.* Prostate-specific antigen (PSA) isoform p2PSA in combination with total PSA and free PSA improves diagnostic accuracy in prostate cancer detection. Eur Urol. 2010; 57:921-7.
45. Le BV, Griffin CR, Loeb S, Carvalhal GF, Kan D, Baumann NA, *et al.* [-2]Proenzyme prostate specific antigen is more accurate than total and free prostate

specific antigen in differentiating prostate cancer from benign disease in a prospective prostate cancer screening study. J Urol. 2010; 183:1355-9.
46. Lazzeri M, Haese A, de la Taille A, Palou Redorta J, McNicholas T, Lughezzani G, *et al.* Serum isoform [-2]proPSA derivatives significantly improve prediction of prostate cancer at initial biopsy in a total PSA range of 2-10 ng/ml: a multicentric European study. Eur Urol. 2013; 63:986-94.
47. Jansen FH, Roobol M, Bangma CH, van Schaik RH. Clinical impact of new prostate-specific antigen WHO standardization on biopsy rates and cancer detection. Clin Chem. 2008; 54:1999-2006.
48. Loeb S, Sokoll LJ, Broyles DL, Bangma CH, van Schaik RH, Klee GG, *et al.* Prospective multicenter evaluation of the Beckman Coulter Prostate Health Index using WHO calibration. J Urol. 2013; 189:1702-6.
49. Loeb S, Partin AW. Review of the literature: PCA3 for prostate cancer risk assessment and prognostication. Rev Urol. 2011; 13:e191-5.
50. Stephan C, Vincendeau S, Houlgatte A, Cammann H, Jung K, Semjonow A. Multicenter evaluation of [-2]proprostate-specific antigen and the prostate health index for detecting prostate cancer. Clin Chem. 2013; 59:306-14.
51. Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. BMJ. 2006; 332:1089-92.
52. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance on undertaking reviews in health care. 2009. Available from: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf. [Accessed March 2014].
53. National Institute for Health and Care Excellence. Diagnostics Assessment Programme manual. 2011. Available from: http://www.nice.org.uk/media/A0B/97/DAPManualFINAL.pdf. [Accessed March 2014].
54. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic test Accuracy Version 100 2010.
55. Jonas DE, Wilt TJ, Taylor BC, Wilkins TM, Matchar DB. Chapter 11: challenges in and principles for conducting systematic reviews of genetic tests used as predictive indicators. J Gen Intern Med. 2012; 27 Suppl 1:S83-93.
56. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, *et al.* The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009; 11:3-14.
57. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155:529-36.
58. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic test Accuracy Version 100 2009.
59. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol. 2006; 6:9.
60. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011; 343.

61. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998; 52:377-84.
62. Hayen A, Macaskill P, Irwig L, Bossuyt P. Appropriate statistical methods are required to assess diagnostic tests for replacement, add-on, and triage. J Clin Epidemiol. 2010; 63:883-91.
63. Takwoingi Y. Meta-analysis of test accuracy studies in Stata: a bivariate approach. Version 1.0. Available from http://srdtacochraneorg/2013.
64. Hendriksen JM, Geersing GJ, Moons KG, de Groot JA. Diagnostic and prognostic prediction models. J Thromb Haemost. 2013; 11 Suppl 1:129-41.
65. Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ economic evaluation working party. BMJ. 1996; 313:275-83.
66. Department of Health. National Schedule of Reference Costs: the main schedule 2012-2013 Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013. [Accessed April 2014].
67. Curtis L. Unit Costs of Health and Social Care 2013. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2013/. [Accessed April 2104].
68. Korfage IJ, Essink-Bot ML, Borsboom GJ, Madalinska JB, Kirkels WJ, Habbema JD, *et al.* Five-year follow-up of health-related quality of life after primary treatment of localized prostate cancer. Int J Cancer. 2005; 116:291-6.
69. Krahn M, Ritvo P, Irvine J, Tomlinson G, Bremner KE, Bezjak A, *et al.* Patient and community preferences for outcomes in prostate cancer: implications for clinical policy. Med Care. 2003; 41:153-64.
70. Shimizu F, Fujino K, Ito YM, Fukuda T, Kawachi Y, Minowada S, *et al.* Factors associated with variation in utility scores among patients with prostate cancer. Value Health. 2008; 11:1190-3.
71. Volk RJ, Cantor SB, Cass AR, Spann SJ, Weller SC, Krahn MD. Preferences of husbands and wives for outcomes of prostate cancer screening and treatment. J Gen Intern Med. 2004; 19:339-48.
72. Scottish Intercollegiate Guidelines Network. Checklist for diagnostic accuracy studies. 2006. Available from: http://www.sign.ac.uk/methodology/checklists.html. [Accessed April 2014].

9 APPENDICES

9.1 Appendix 1: Section on modelling possibilities from final NICE scope

4.2 Modelling possibilities
There are a number of possible pathway options for the economic model.
These include:

PCA3 or PHI testing is carried out after an initial negative biopsy to determine, in combination with information on other risk factors, if a repeat biopsy is needed followed by standard care as described in CG175.

PCA3 or PHI testing is carried out after an initial negative biopsy to determine, in combination with information on other risk factors, if an mpMRI is needed. If the mpMRI is positive, a biopsy would be performed followed by standard care as described in CG175.

PCA3 or PHI testing is carried out after both an initial negative biopsy and a negative mpMRI to determine, in combination with information on other risk factors, if a repeat biopsy is needed followed by standard care as described in CG175.

In all cases the impact of performing either a template, saturation, or a targeted biopsy when undertaking a second biopsy should be considered.

The impact of potentially increasing the detection of clinically insignificant prostate cancer by the use of additional diagnostic tests should also be considered. Conversely, the impact of potentially decreasing overtreatment of prostate cancer through improved risk stratification by the use of additional diagnostic tests should also be considered.

9.2 Appendix 2: Draft search strategy

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)

Search for studies on PCA3 or phi

1. exp prostatic neoplasms/
2. (prostat* adj3 (cancer or carcinoma* or neoplasm* or malignant* or tumor* or tumour*)).tw.
3. or/1-2
4. (Prostat* adj2 cancer* adj2 (antigen* or gene*) adj2 "3").tw.
5. (PCA3 or PCA-3 or "PCA 3").tw.
6. uPM3.tw.
7. ("differential display code 3 antigen" or DD3).tw.
8. progensa.tw.
9. or/4-8
10. prostate health index.tw.
11. Beckman Coulter.tw.
12. (proPSA or p2proPSA).tw.
13. or/10-12
14. or/9,13
15. 3 and 14
16. exp animals/ not humans/
17. nonhuman/ not human/
18. or/16-17
19. 15 not 18
20. limit 19 to yr=2000-2014

Search for reviews on comparators

1. exp Magnetic Resonance Spectroscopy/
2. magnetic resonance imaging/ or exp diffusion magnetic resonance imaging/
3. magnetic resonance imag$.tw.
4. magnetic resonance spectroscop*.tw.
5. mrs.tw.
6. (dynamic contrast enhanced adj3 (MRI or magnetic)).tw.
7. dce-mri.tw.
8. (diffusion weight$ adj3 (MRI or magnetic)).tw.

9. dw-mri.tw.

10. ((multi-parametric or multiparametric or mp) adj (MRI or magnetic).tw.

11. or/1-10

12. exp Prostate/ah, pa, us [Anatomy & Histology, Pathology, Ultrasonography]

13. (transrectal adj (biops* or ultrasound or ultrason*)).tw.

14. trus.tw.

15. exp Biopsy, Needle/

16. (biopsy or biopsies or pathol* or histopathol*).tw.

17. or/12-16

18. exp Prostate-Specific Antigen/

19. psa.tw.

20. prostat* specific antigen*.tw.

21. or/18-20

22. exp nomograms/

23. nomogram*.tw.

24. (neural adj2 network).tw.

25. or/22-24

26. exp prostatic neoplasms/

27. (prostat* adj3 (cancer or carcinoma* or neoplasm* or malignant* or tumor* or tumour*)).tw.

28. or/26-27

29. or/11,17,21,25

30. 28 and 29

31. exp meta-analysis/

32. exp Meta-Analysis as Topic/

33. meta-analys*.mp. or (meta adj analys*).ti,ab.

34. meta-regress*.mp. or (meta adj regress*).ti,ab.

35. meta analysis.pt.

36. systematic review.ti.

37. or/31-36

38. 30 and 37

9.3 Appendix 3: Draft data extraction forms

The data to be extracted will include, but not be limited to:

1. **Review of analytic validity**
   - study design
   - type of report (peer-reviewed or unpublished data)
   - details of funding
   - source population
   - specific methods/platforms evaluated
   - pre-analytic variables studied
   - number of samples tested
   - timing and locations of repeat assays
   - results reported, including means and coefficient of variation, test failure rate

2. **Review of clinical validity**

Study characteristics
- type of report (abstract, full manuscript, interim report)
- type of study
- methodological details of study
- location of study
- method of allocation to intervention or comparator test
- details of funding

Participants
- selection criteria
- number in study
- age
- ethnicity

Investigation before first biopsy

- tests performed before initial biopsy, PSA levels
- type of first biopsy
- number of cores taken
- definition of negative results

Intervention tests

- details of tests used
- timing of test
- thresholds used
- test pathway description

Comparator tests

- details of tests used
- timing of test
- thresholds used

Reference standard

- type of second biopsy
- use of MRI targeting technology
- number of cores sampled
- timing of biopsy
- definition of positive biopsy
- histopathology procedures and expertise
- differences in reference standard depending on intervention or comparator test results

Results

The available data on all diagnostic outcomes will be recorded. If multiple thresholds are used and reported, outcomes will be recorded for each threshold.

- estimates of the intervention or comparator test (Means (SD), proportion positive) in men with positive and negative results on second biopsy
- specificity and sensitivity for different cut-off points of PCA3, phi or PSA
- comparison of AUC for different tests or test combinations
- gain in sensitivity and specificity estimates by adding intervention test as derived from ROC curves
- differences in Gleason score or stage of cancers detected
- results of logistic regression analyses

3. Review of clinical utility

Data on *study characteristics, participants* and *investigation before first biopsy* will be recorded as for review of clinical validity.

The available data on all clinical outcomes reported will be recorded including:

- morbidity and mortality from biopsies
- morbidity and mortality from treatment of diagnosed cancer
- health related quality of life
- patient reported outcomes

4. Details of economic data extraction

Cost effectiveness data extraction will include, but may not be limited to:

Study characteristics

- type of evaluation and synthesis
- intervention
- study population/disease
- time period of study

Cost data and cost data sources

- cost items
- cost data sources
- country, currency year

Outcome data and data sources

- range of outcomes
- efficiency data sources
- modelling method and data sources
- probabilities and assumptions of models

Cost effectiveness

- cost-effectiveness ratios
- subgroup analysis and results
- sensitivity analysis and results
- authors conclusion

9.4 Appendix 4:QUADAS-2 assessment form

SIGN	Methodology Checklist 5: Studies of Diagnostic Accuracy[57,72] *This checklist is based on the work of the QUADAS-2 team at Bristol University (http://www.bris.ac.uk/quadas/).*
Study identification (*Include author, title, reference, year of publication*)	
Guideline topic:	Key Question No:

Before completing this checklist, consider:

1. Is the paper really a study of diagnostic accuracy? It should be comparing a specific diagnostic test against another, and **not** a general paper or comment on diagnosis.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist..

Reason for rejection: Reason for rejection: 1. Paper not relevant to key question □
2. Other reason □ (please specify):

Checklist completed by:

All the questions in the following sections have associated footnotes providing short explanations behind each of the questions. Users who want more detailed explanations should consult the QUADAS-2: Background Document.

DOMAIN 1 – PATIENT SELECTION

Risk of bias

	In a well conducted diagnostic study…	***Is that true in this study?***
1.1	A consecutive sequence or random selection of patients is enrolled.[i]	Yes ☐ Can't say ☐ No ☐
1.2	Case – control methods are not used.[ii]	Yes ☐ Can't say ☐ No ☐
1.3	Inappropriate exclusions are avoided.[iii]	Yes ☐ Can't say ☐ No ☐
Applicability		
1.4	The included patients and settings match the key question.[iv]	Yes ☐ Can't say ☐ No ☐

DOMAIN 2 – INDEX TEST		
Risk of bias		
In a well conducted diagnostic study…		***Is that true in this study?***
2.1	The index test results interpreted without knowledge of the results of the reference standard.[v]	Yes ☐ Can't say ☐ No ☐
2.2	If a threshold is used, it is pre-specified.[vi]	Yes ☐ Can't say ☐ No ☐

i.Studies should enrol either all eligible patients suspected of having the target condition during a specified period, or a random sample of those patients. The essential point is that investigators should have no freedom of choice as to which individual patients are or are not included.
ii. There is evidence that studies comparing patients with known disease with a control group without the condition tend to exaggerate diagnostic accuracy.
iii.Inappropriate exclusions may result in either overestimates (eg by excluding 'difficult to diagnose' patients) or underestimates (eg by excluding patients with 'red flags' suggesting presence of disease) of the degree of diagnostic accuracy.
iv. Patients included in the study should match the target population of the guideline in terms of severity of the target condition, demographic features, presence of differential diagnosis or co-morbidity, setting of the study and previous testing protocols.
v. This is similar to the question of 'blinding' in intervention studies. The index test should always been done first, or by a separate investigator with no knowledge of the outcome of the reference test.
vi. Bias can be introduced if a threshold level is set after data has been collected. Any minimum threshold should be specified at the start of the trial.