CONFIDENTIAL UNTIL PUBLISHED

External Assessment Group's Report
MRI fusion biopsy in people with suspected prostate cancer

Produced by	CRD and CHE Technology Assessment Group, University of York, [redacted]
Authors	Alexis Llewellyn, Research Fellow, CRD Thai Han Phung, Research Fellow, CHE Marta Soares, Professor, CHE Lucy Beresford, Research Fellow, CRD David Glynn, Research Fellow, CHE Melissa Harden, Senior Information Specialist, CRD Ruth Walker, Research Fellow, CRD Ana Duarte, Research Fellow, CHE Sofia Dias, Professor of Health Technology Assessment, CRD
Correspondence to	Professor Sofia Dias, CRD, University of York[redacted]
Date completed	02/11/2022

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135477.

Declared competing interests of the authors

'None'.

Acknowledgements

We would like to thank specialist DAC members Dr Oliver Hulson, Consultant Radiologist at Leeds Teaching Hospital Trust, Ms Lillian White, Clinical Nurse Specialist at NHS Ayshire and Arran, and Mr Hide Yamamoto, Consultant Urological Surgeon at Maidstone and Tunbridge Wells Trust for providing helpful advice during the project. We also thank Ms Lillian White for commenting on an early draft of the report. We would also like to thank Prof Hashim Ahmed for providing helpful advice during this project; Mr Connor Evans, Research Trainee at CRD, for helping to edit the final version of this report; and Prof Stephen Palmer for advice on the conceptualisation of the economic model.

We thank Dr Bhash Naidoo and the NICE Centre for Guidelines who developed the NICE NG131 health economics model for sharing it. The review of the electronic version of the model contributed to inform the York model conceptualisation. Similarly, we would like to thank Dr Joanne Lord and the Southampton TAR group who developed the Southampton model for sharing it. The electronic version of the Southampton model was used to support both model conceptualisation and parameterisation of the York model.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Llewellyn A, Phung TH, Soares M, Beresford L, Glynn D, Harden M, Walker R, Duarte A, Dias S. MRI fusion biopsy in people with suspected prostate cancer: A Diagnostics Assessment Report. York EAG, 2022.

Contributions of authors

Lucy Beresford (Systematic Reviewer) contributed to the protocol, performed the systematic review and wrote parts of the clinical sections.

Sofia Dias (Statistician) contributed to the protocol, adapted the network meta-analysis model and performed the meta-analyses, wrote up the meta-analysis section results, and oversaw the conduct and writing of the clinical effectiveness sections and the report as a whole. Sofia has overall responsibility for the clinical effectiveness sections of the report.

Ana Duarte (Health Economist) developed the economic model, conducted analyses over this model and validated it. Ana had overall responsibility for the cost-effectiveness sections of the report. She wrote the cost-effectiveness sections of the protocol and of the report.

David Glynn (Health Economist) reviewed the prostate cancer long-term outcome evidence and developed the inference sub-model to inform the economic model.

Melissa Harden (Information Specialist) designed and ran all searches for the project, managed the library of references, and wrote the search sections of the report.

Alexis Llewellyn (Systematic Reviewer) contributed to the protocol, performed the systematic review, wrote the background and most of the sections on clinical effectiveness.

Thai Han Phung (Health Economist) assisted with the economic evidence reviews and with the development of the economic model, its analyses and validation, and contributed to the writing of the cost-effectiveness sections.

Marta Soares (Health Economist) developed the methods to extend the evidence synthesis for use in the economic model and conducted the required analyses over the extended synthesis model. She provided oversight and contributed to the development of the inference sub-model to inform the

economic model. She wrote the cost-effectiveness sections of the protocol and main report, and provided leadership support to the economic sections.
Ruth Walker (Systematic Reviewer) contributed to the protocol and backgound materials.

Note on the text

No commercial-in-confidence (CIC) or academic-in-confidence (AIC) data are included in this report.

Copyright statement

ABSTRACT

Background

MRI localizes cancer in the prostate, allowing targeted biopsy with or without transrectal ultrasound-guided systematic biopsy. Targeted biopsy methods include cognitive fusion, where prostate lesions suspicious on MRI are targeted visually during live ultrasound, and software fusion, where computer software overlays the MRI image onto the ultrasound in real time. The effectiveness and cost-effectiveness of software fusion technologies compared with cognitive fusion biopsy are uncertain.

Objectives

To assess the clinical and cost-effectiveness of software fusion biopsy technologies in people with suspected localised and locally advanced prostate cancer.

A systematic review was conducted to evaluate the diagnostic accuracy, clinical efficacy and practical implementation of nine software fusion devices compared to cognitive fusion biopsies, and with each other, in people with suspected localised, or locally-advanced prostate cancer. Comprehensive searches including MEDLINE, and Embase were conducted up to August 2022 to identify studies which compared software fusion and cognitive fusion biopsies in people with suspected prostate cancer. Risk of bias was assessed with QUADAS-C.

A network-meta analysis comparing software fusion and cognitive fusion with or without concomitant systematic biopsy, and systematic biopsy alone was conducted. Additional outcomes, including safety and usability, were synthesised narratively.

A *de novo* decision model was developed to estimate the cost-effectiveness of targeted software fusion biopsy relative to cognitive fusion biopsy with or without concomitant systematic biopsy for prostate cancer identification in biopsy naïve people . Scenario analyses were undertaken to explore the robustness of the results to variation in the model data sources and alternative assumptions.

Results

23 studies (3773 patients with software fusion, 2154 cognitive fusion) were included. Evidence was available for seven of the nine fusion devices specified in the protocol and at high risk of bias.

Results suggest that patients undergoing cognitive biopsy may show: i) a higher probability of being classified as not having cancer, ii) similar probability of being classified as having non-clinically significant cancer (ISUP grade 1), and iii) lower probability of being classified at higher ISUP grades, particularly ISUP 2. Similar results were obtained when comparing between same biopsy methods where both were combined with systematic biopsy. Evidence was insufficient to conclude whether

any individual devices were superior to cognitive fusion, or whether some software fusion technologies were superior to others.

Uncertainty in the relative diagnostic accuracy of software fusion vs. cognitive fusion reduce the strength of any statements on its cost-effectiveness. So although the economic analysis suggests ICERs for software fusion biopsy vs. cognitive fusion of £1,826 and £5,623 per additional QALY with or with concomitant systematic biopsy, respectively.

Limitations

There was insufficient evidence to explore the impact of effect modifiers.

Conclusions

Software fusion biopsies may be associated with increased cancer detection in relation to cognitive fusion biopsies, but the evidence is at high risk of bias. Sufficiently powered, high-quality studies are required. Cost-effectiveness results should be interpreted with caution given the limitations of the diagnostic accuracy evidence.

479 words

1 SCIENTIFIC SUMMARY

1.1 Background

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

1.2 Objectives

This project aimed to assess the clinical and cost-effectiveness of software fusion biopsy systems in people with suspected localised and locally advanced prostate cancer.

1.3 Methods

1.3.1 Systematic review

A systematic review of the diagnostic accuracy, clinical effectiveness, safety and practical implementation of nine software fusion systems compared with cognitive fusion and with each other, in people suspected prostate cancer according to MRI was conducted.

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

Studies of people with suspected prostate cancer who have had an MRI scan that indicates a significant lesion (Likert or PI-RADS score of 3 or more), including biopsy naïve and repeat biopsy patients with a previous negative prostate biopsy, and comparing software fusion with cognitive fusion or with another software fusion device, were included. The following software fusion technologies were included: Artemis (InnoMedicus Artemis), Biojet (Healthcare Supply Solutions Ltd), BiopSee (Medcom), bkFusion (BK Medical UK Ltd and MIM Software Inc), Fusion Bx 2.0

(Focal Healthcare), FusionVu (Exact Imaging), iSR'obot Mona LisaTM (Biobot iSR'obot), KOELIS Trinity (KOELIS and Kebomed) and UroNav Fusion Biopsy System (Phillips). Previous versions were also eligible. In-bore (or in-gantry) biopsies were excluded. Prospective, randomised and non-randomised comparative studies were included, and retrospective evidence where no prospective evidence could be found for an eligible software fusion device. To provide sufficient evidence for a network meta-analysis, within-patient comparisons or RCTs between software fusion and systematic biopsy, and between cognitive fusion and systematic biopsy, were also eligible to inform indirect comparisons of diagnostic accuracy.

Two researchers independently screened the titles and abstracts of all reports identified by the bibliographic searches and of all full-text papers subsequently obtained. Data extraction and quality assessment were conducted by at least one researcher and checked by a second. Risk of bias of diagnostic accuracy studies was assessed using QUADAS-C.

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

1.3.2 Economic analysis

Cost-effectiveness evidence comparing software fusion biopsy systems with cognitive fusion for targeted prostate biopsy in men with suspected prostate cancer was identified by the abovementioned searches, with evidence narratively summarised and tabulated. Studies were appraised for their quality, generalisability and appropriateness to inform the decision problem as defined by the NICE DAR scope. A targeted search was conducted to identify evidence to support the development of a de novo decision model. The searches aimed to identify cost-effectiveness evidence of diagnostic strategies at the point of biopsy to support the model conceptualisation. Evidence was reviewed to i) identify value components of the biopsy approaches, ii) characterise alternative mechanisms of evidence linkage from disease prevalence, diagnostic accuracy, choice of treatment to final outcomes, and iii) identify any UK relevant sources of evidence.A *de novo* decision analytic model was developed to estimate the cost-effectiveness of software fusion compared to cognitive fusion. The model evaluated two strategies for two alternative comparisons: i) targeted software fusion biopsy vs. targeted cognitive biopsy and ii) combined (targeted and systematic) software fusion biopsy vs. combined cognitive biopsy. The four strategies could not be incrementally compared due to the

mechanism of evidence generation for the diagnostic accuracy, which relied on separate evidence networks.

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

The model required the development of i) an extension to the evidence synthesis to allow quantifying the extension of test misclassification in the diagnostic model with software fusion biopsy and cognitive fusion biopsy, and ii) an inference model to derive unobservable transition probabilities for the long-term model.

1.4 Results

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

1.4.1 Diagnostic accuracy

Across all analyses results must be interpreted with caution due to the high risk of bias in the evidence base and wide uncertainty over the results. Results suggest that patients undergoing cognitive biopsy may show: i) a higher probability of being classified as not having cancer, ii) similar probability of being classified as having non-clinically significant cancer (ISUP grade 1), and iii) lower probability of being classified at higher ISUP grades, particularly ISUP 2. Similar results were obtained when comparing between same biopsy methods where both were combined with systematic biopsy.

Additional meta-analyses of cancer detection rates suggest that, compared with cognitive fusion biopsy, software fusion may identify more prostate cancer (any grade) (OR 1.30; 95% CrI 1.06, 1.61) and more non-clinically significant cancer (ISUP 1) (OR 1.98; 95% CrI 1.28, 3.06). Adding systematic biopsy to cognitive or software fusion may increase the detection of all prostate cancer and of clinically significant cancer, and from this evidence there is no suggestion that software fusion with concomitant systematic biopsy is superior to cognitive fusion with systematic biopsy.

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

No evidence was found for other outcomes.

1.4.2 Clinical effectiveness

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

1.4.3 Cost-effectiveness

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

The base-case cost-effectiveness analysis suggests for the targeted biopsy and the combined biopsy comparisons, that software fusion strategy is on average costlier and yields greater QALYs than the cognitive fusion strategy, resulting in a deterministic ICER of £5,623 and £1,826 per additional QALY for each comparison, respectively. These ICERs are below the lower bound of the cost-effectiveness threshold range recommended by NICE, suggesting that software fusion may be cost-effective compared to cognitive fusions in both the targetd and the combined comparisons. However, these results should be interpreted cautiously given the uncertainties in the relative diagnostic accuracy evidence which informs the model. The probabilistic analysis suggests a higher probability of cost-effectiveness for software fusion vs. cognitive fusion at the range of cost-effectiveness thresholds recommended by NICE (0.64 and 0.68 at £20,000 and £30,000 per additional QALY for targeted software fusion biopsy).

1.5 Discussion

This assessment includes a broad, comprehensive literature search for software and cognitive fusion technologies and has been conducted following recognised guidelines to ensure high quality. The

review identified evidence on the diagnostic accuracy of nine software fusion technologies, and is the first systematic review to formally compare the relative accuracy of software fusion and cognitive fusion, with and without systematic biopsy, as well as different software fusion devices, using both direct and indirect evidence in a network meta-analysis. Unlike recent systematic review evidence, our review found that software fusion increased detection of clinically insignificant cancer compared with cognitive fusion.

Our review has a number of limitations. The evidence included in the systematic review is at high risk of bias overall. There was variation in patient and study characteristics. Biopsy naïve patients, who form the large majority of patients eligible for targeted biopsy, were underrepresented, although there was insufficient evidence to evaluate whether the relative accuracy of software and cognitive fusion differed between biopsy naïve and repeat biopsy patients. There was insufficient evidence to explore the impact of a number of other potential effect modifiers, including lesion location, operator experience, biopsy routes and anaesthesia methods. There were few studies per comparison, not all studies reported outcomes by all cancer grades, and most estimates from the meta-analyses were imprecise, particularly at higher cancer grades where data was most sparse. The network meta-analyses relied on the assumption that cognitive fusion was equivalent across different centres, which is uncertain.

1.6 Conclusions

Software fusion biopsies identify more clinically insignificant cancer than cognitive fusion biopsies, although there is no evidence that software fusion detects more clinically significant cancer. Both software fusion and cognitive fusion biopsy miss clinically significant cancer lesions, and the addition of standard-systematic biopsy increases the detection of all prostate cancer and clinically significant cancer. There is insufficient evidence to conclude on the relative accuracy and clinical effectiveness of different software devices.

1.6.1 Recommendations for further research

High-quality, sufficiently powered randomised controlled trial evidence comparing software fusion biopsy with cognitive fusion biopsy is required to address limitations from the existing evidence. Improved reporting of diagnostic accuracy outcomes would enable future syntheses to make use of a larger body of evidence.

1.6.2 Study registration

The protocol for this review is registered on PROSPERO (CRD42022329259).

PLAIN ENGLISH SUMMARY

Men with an MRI scan that shows possible prostate cancer are offered prostate biopsies, where samples of the prostate tissue are collected with a needle, to confirm the presence and severity of cancer. Different biopsy methods exist. In a cognitive fusion biopsy, clinicians will target abnormal looking parts of the prostate by looking at the MRI scan alongside 'live' ultrasound images. During a software fusion biopsy, a computer software is used to overlay the MRI scan onto the ultrasound image. This project evaluated whether software fusion is better at detecting cancer compared with cognitive fusion biopsy, and whether it represents value for money for the NHS.

A comprehensive review of the literature on software fusion technologies was performed. Data were combined and re-analysed, and the quality of the evidence was assessed. Economic analyses investigated whether software fusion biopsies are sufficient value for money.

We found that software fusion biopsies detect more low-grade, insignificant disease than cognitive fusion biopsies, although there is no evidence that software fusion identified more significant cancer that would require treatment. We found no evidence that any software fusion devices were superior to others. Using additional, random biopsies alongside software or cognitive fusion would increase the detection of prostate cancer.

This project also looked for evidence on the value for money of the software fusion biopsies to detect prostate cancer and found no relevant studies. We weighed the costs and the benefits of software fusion biopsy compared to cognitive fusion to determine whether it could be a good use of NHS money. The poor quality of information to support our analysis, makes consideration on the value of the technologies largely unknown.

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

ADT	Androgen deprivation therapy
AE	Adverse event
AEs	Adverse events
AIC	Academic in confidence
ASG	Assessment subgroup
bpMRI	Bi-parametric magnetic resonance imaging
CF	Cognitive fusion
CI	Confidence interval
CIC	Commercial in confidence
CNS	Clinically non-significant
CPG	Cambridge Prognostic Group
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Clinically significant
CsPCa	Clinically significant prostate cancer
DAR	Diagnostics Assessment Report
DCD	Diagnostics Consultation Document
DRE	Digital rectal examination
DTX	Docetaxel
EAG	External Assessment Group
EAU	European Association of Urology
FAD	Final Appraisal Document
FN	False negative
GATP	General anaesthesia transperineal
GG	Grade Group
GIN	Guidelines International Network
GP	General practitioner
GS	Gleason score
HR	Hazard ratio
HRG	Healthcare resource groups
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
INHB	Incremental net health benefit
IPD	Individual participant data
IQR	Interquartile range

ISUP	International Society of Urological Pathology
IT	Information technology
LATP	Local anaesthesia transperineal
LATRUS	Local anaesthesia transrectal ultrasound
LHRH	Luteinizing hormone-releasing hormone
mpMRI	Multiparametric magnetic resonance imaging
MRI	Magnetic resonance imaging
NA	Not applicable
NC	No cancer
NG131	NICE Guideline 131
NHB	Net health benefit
NHS	National Health Service
NHSCII	NHS Cost Inflation Index
NICE	National Institute for Health and Care Excellence
NICE	The National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
NMA	Network meta-analysis
NPCA	National Prostate Cancer Audit
NPV	Negative predictive value
NR	Not reported
NRFT	No recurrence following treatment
OS	Overall survival
PACS	Picture archiving and communication system
PADT	Primary androgen deprivation therapy
PCa	Prostate cancer
PFS	Progression free survival
PI-RADS	Prostate imaging - reporting and data system
PRFT	Possible recurrence following treatment
PSA	Prostate specific antigen
PSS	Personal Social Service
QADAS	Quality Assessment tool of Diagnostic Accuracy Studies
QALE	Quality-adjusted life expectancy
QALYs	Quality-adjusted life years
QoL	Quality of life
RCT	Randomised controlled trial
RFI	Response to information request
RR	Relative risk/risk ratio
SAE	Serious adverse event
SCM	Specialist Committee Member

SD	Standard deviation
SE	Standard error
SF	Software fusion
SOC	Standard of care
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TN	True negative
TP	Transperineal biopsys
TRUS	Transrectal ultrasound
TSB	Template-guided saturation biopsy
TSD	Technical Support Document
TTMB	Template-guided mapping biopsy
UK	United Kingdom
US	United States
UTI	Urinary tract infection
UTIs	Urinary tract infections
VAT	Value-added tax

Glossary

Preferred term	**Definition**
Targeted biopsy	Biopsy where the site (or sites) for sampling is (or are) targeted based on the location of one or more potentially cancerous lesions identified by an MRI scan. Includes software fusion biopsy, cognitive fusion biopsy, and in-bore biopsy. Also referred to as MRI-targeted.
Software fusion biopsy	Software fusion is software-based technology used to fuse pre-biopsy MRI image and real-time ultrasound images to create a detailed 3D image. Software fusion biopsy refers to biopsies where software fusion is used to guide and record the placement of biopsy needles. Also referred to as MRI fusion.
Cognitive fusion biopsy	When the operator views both sets of MRI and ultrasound images and mentally translates the MRI target lesions onto the real-time ultrasound images during the biopsy procedure, to guide the placement of biopsy needles. Also referred to as visual estimation or visual registration.
In-bore biopsy	Technique that involves performing the prostate biopsy in the MRI scanner, where the needle is inserted within the MRI machine, and placement is guided by the MRI images in real time. Also referred to as in-gantry biopsy.
Systematic biopsy	Biopsy method where samples are taken in a systematic fashion from different regions of the prostate according to a predefined scheme. The number of cores sampled can range from six to 14, and is most commonly 12. Also referred to as random biopsy or 12-core biopsy.
Template biopsy	Biopsy method where samples are taken in a systematic fashion from different regions of the prostate using a grid template. The minimum number of cores is typically 20. Also referred to as template prostate mapping.
Route of access	Route employed to reach the prostate with a biopsy needle. Can be either via the rectum (transrectal) or the perineum (transperineal). Also referred to as biopsy route
Transrectal ultrasound (TRUS)-biopsy	Where a biopsy needle is inserted through the rectal wall via the anus, and positioning is informed by ultrasound imaging.

Active surveillance	Monitoring of a person following a diagnosis of prostate cancer, with a view to switching to radical treatment if the cancer progresses. Aims to prevent the risk of overtreatment by avoiding immediate radical intervention. Active surveillance typically includes regular monitoring of prostate-specific antigen (PSA) levels and digital rectal examination.
Gleason system	A system used to grade prostate cancer cells to estimate how quickly they are likely to grow (Gleason grade). Grade Group 1 is the least aggressive, indicating that the cancer is likely to grow very slowly, if at all. Grade Group 5 is the most aggressive, indicating the cells look very abnormal and the cancer is likely to grow quickly. Since prostate tumours are often made up of cancerous cells that have different grades, two grades are assigned for each patient. A primary grade is given to describe the cells that make up the largest area of the tumour and a secondary grade is given to describe the cells of the next largest area. For example, a Gleason Score written as 3+4=7 indicates that most of the tumour is grade 3 and the next largest section of the tumour is grade 4. To help with outcome prediction and patient communication, Gleason scores ≤6, 3+4, 4+3, 8 and 9-10, respectively, can be reported as five risk groups defined by the International Society of Urological Pathology (ISUP), i.e. ISUP grades 1-5.
ISUP Gleason grades	Grouping of Gleason scores into risk groups defined by the International Society of Urological Pathology (ISUP) to help with outcome prediction and patient communication.
Likert score	A Likert score is reported using a 5-point Likert scale. The Likert scale, when used in the diagnosis of prostate cancer, takes into account clinical factors and lesion size on MRI. A score of 1 indicates prostate cancer is very unlikely and a score of 5 indicates prostate cancer is very likely. Likert scores are used to help decide whether or not to have a prostate biopsy at the current time. The Likert score differs from the PI-RADS score in that it takes into account clinical factors and does not require the MRI to be conducted in a particular sequence.
PI-RADS score	The Prostate Imaging Reporting and Data System (PI-RADS) score is a system whereby each lesion identified by MRI is assigned a score from 1 to 5 to indicate the likelihood of clinically significant cancer (where 1 is very low and 5 is very high). PI-RADS v2 is the current validated version. It differs from the Likert score in that it does not take into account clinical factors and it requires the MRI to be conducted in a particular order.
Watchful waiting	Monitoring of a person diagnosed with prostate cancer where any potential treatment offered aims to control rather than cure the prostate cancer (palliative rather than curative intent).
Grid and stepping device	A stepping device is used in prostate biopsy to cradle the ultrasound probe. On this device, a grid can be attached. A grid (or template) is used in transperineal biopsies. The grid, which is placed in front of the perineum, includes a number of holes in which the biopsy needle can be inserted. Each hole is correlated to numbers and letters which allow for precise sampling of prostate. Also referred to as a template (the grid) and a stepper (stepping device)
Freehand	A biopsy in which the ultrasound probe is held in the hand, rather than being supported by a stepping device. This allows the probe to be moved in all directions. A needle attached to the ultrasound probe is then used to puncture the perineum before the biopsy needle is passed through. The biopsy needle can be pivoted to take the samples, reducing the number of puncture sites on the perineum.
Double freehand	A transperineal biopsy technique whereby the ultrasound probe is held in the hand, rather than being supported by a stepping device. Unlike the freehand technique, the introducer needle is not attached to the ultrasound probe and is held in the other hand.
Rigid registration	During software fusion with rigid registration, the MRI image is fixed, and is not adjusted to match the ultrasound image when potential deformation to the prostate that may occur during the biopsy.
Elastic registration	During software fusion with elastic registration, the MRI image is altered to match the ultrasound image, to adjust for potential deformation to the prostate during the biopsy. Also referred to as non-rigid registration.
Semi-robotic arm	Used in prostate biopsies, the semi-robotic arm is attached to the ultrasound probe. It allows the operator to manoeuvre the probe into the position of interest whilst ensuring a consistent level of pressure on the prostate to reduce prostate deformation.

ASSESSMENT GROUP REPORT

2 BACKGROUND AND DEFINITION OF THE DECISION PROBLEM

2.1 Description of health problem

Prostate cancer (PCa) is the most commonly diagnosed cancer in men in the UK; it accounts for more than a quarter (27%) of all male cancer diagnoses in 2016-2018.Cancer Research UK [1] It is the second most common cause of cancer death in males in the UK, accounting for 14% of all cancer deaths. The estimated lifetime risk of a PCa diagnosis is one in eight for males born in the UK.[2, 3] Over 57,000 new cases were diagnosed in 2018, with an estimated ten-year survival rate of 77.6%. Since the early 1990s, estimates of PCa incidence rates have increased by nearly half (48%) in males in the UK (2016-2018) and are projected to rise by 12% between 2014 and 2035, to 233 cases per 100,000 males by 2035.[3]

Early stage diagnosis is associated with improved survival outcomes compared with patients diagnosed at the latest stage of the disease. Prostate cancer primarily affects people aged 50 years or more, and the risk of developing PCa increases with age.[3] In England and Wales, 87% of people diagnosed with PCa are aged 60 years or older, [4] and on average each year around a third of new cases (34%) were in males aged 75 and over.[2] People of African family background and individuals with a family history of PCa are at higher risk of PCa.[5, 6]

Prostate cancer might be suspected if any of the following symptoms cannot be attributed to other health conditions: lower urinary tract symptoms, such as frequency, urgency, hesitancy, terminal dribbling and/or overactive bladder; erectile dysfunction; haematuria; lower back or bone pain; lethargy, and weight loss.

The descriptor ‘clinically significant’ is widely used to differentiate PCa that may lead to morbidity or death from types of PCa that do not. This distinction is important as insignificant PCa that does not cause harm is common.[7] Autopsy studies in men who died of causes other than PCa indicate that there is a significant prevalence of non-clinically significant prostate in the general male population, which increases with age.[7] Prostate cancer screening may therefore lead to overdiagnosis, by identifying cancers that are not destined to cause morbidity or mortality. Men with these cancers are at risk of being harmed by early detection and unnecessary treatment,[8, 9] such as radical prostatectomy or radiotherapy with no additional mortality benefit compared to an active surveillance approach, which includes regular monitoring of prostate-specific antigen (PSA) levels and digital rectal examination. On the other hand, individuals with undetected cancer or with lesions incorrectly classed as benign may miss out on relevant treatment. Clinical guidelines have focused efforts to address the risk of

overtreatment and undertreatment of PCa, notably with recent updates to diagnosis pathways and refinements to risk stratification of cancer lesions.[10-12]

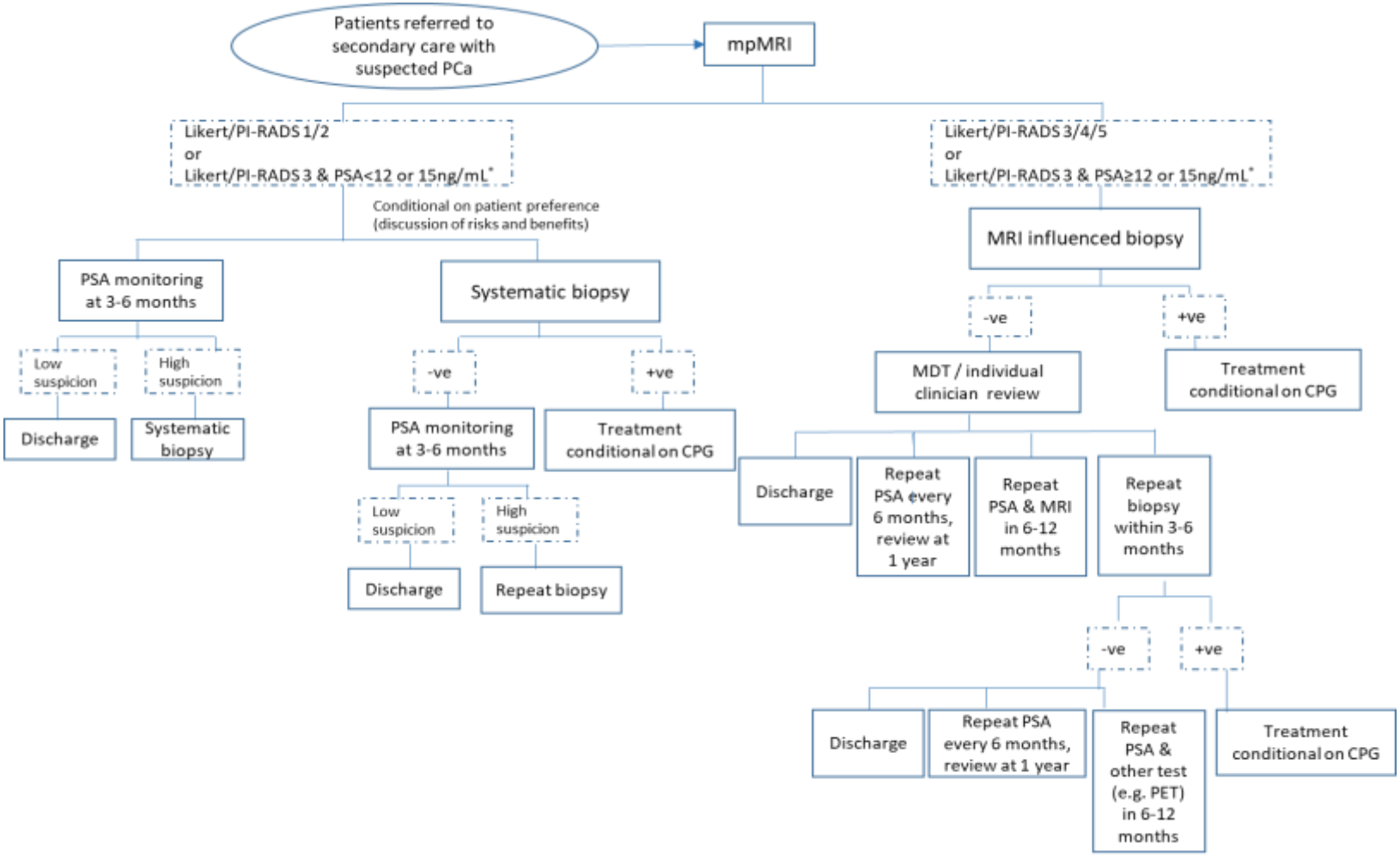
2.2 Care pathways for the diagnosis and management of prostate cancer

2.2.1 Referral to suspected cancer pathway

There is no screening programme in the UK for PCa, although PSA testing is available for asymptomatic individuals above 50 requesting this test.[13] For people presenting to primary care with certain clinical signs and symptoms that may indicate PCa, NICE's guideline for suspected cancer recognition and referral advises to consider a PSA test and digital rectal examination to assess for PCa in men with: any lower urinary tract symptoms (such as nocturia, urinary frequency, hesitancy, urgency or retention) or erectile dysfunction or visible haematuria.[14] The guideline recommends men should be referred using a suspected cancer pathway (for an appointment within two weeks) for PCa if their PSA levels are above the age-specific reference range or if their prostate feels malignant (hard, or lumpy) on digital rectal examination. The NHS Faster Diagnosis Standard requires that patients are diagnosed or have cancer ruled out within 28 days of being referred urgently by their general practitioner (GP) for suspected cancer,[15] and NICE requires that GPs should have direct access to appropriate imaging tests.[16]

Figure 1 summarises the EAG's interpretation of the pathway for the diagnosis and care of individuals with suspected PCa according to NICE guidance (NG) 131 and the NHS timed PCa pathway, which was validated by clinical advisers to the EAG.[12, 17]

Figure 1 Diagnostic and care pathway for individuals with suspected prostate cancer



*per mL of prostate volume; CPG, Cambridge Prognostic Group; MDT, multidisciplinary team.

2.2.2 MRI for suspected cancer

NICE's guideline for diagnosis and management of PCa advises that, in patients with suspected clinically localised PCa, multiparametric magnetic resonance imaging (mpMRI) should be offered as the first-line investigation, but not to those patients who would not be able to have radical treatment.[12] This guidance superseded prior guidance which recommended transrectal ultrasound (TRUS)-guided systematic biopsy as first-line test. Introduced in the 2019 review of the guidelines, the recommendation to offer first-line mpMRI followed the results of PROMIS and PRECISION studies which found a greater negative predictive value (NPV) with mpMRI as first-line diagnostic test compared with the traditional standard-of-care use of TRUS-guided systematic biopsy.[18, 19]

The results of the MRI can be reported using a 5-point Likert scale as recommended in NG131, which estimates the risk that an area seen on the MRI scan may be a cancer or not. The Prostate Imaging Reporting and Data System (PI-RADS) is an alternative to the Likert scale assessment of MRI results.[20-22] Here, each lesion is assigned a score from 1 to 5, with higher scores, usually PI-RADS 4 and 5, indicating a higher likelihood of clinically significant cancer.

2.2.2.1 mpMRI and compliance with NICE guidance

Uptake of MRI prior to biopsy in England and Wales has significantly increased in recent years, from 37% in 2017 to 87% in 2019. Data from 10 of 14 trusts in Scotland also indicate that uptake of a pre-biopsy bi-parametric MRI (bpMRI) or mpMRI as first-line diagnostic ranged from 75% to 100% across centres, although most trusts have not yet met the new NHS Scotland 95% target.[23, 24] TRUS biopsy is still offered as first-line investigation for some patients, although the practice is becoming increasingly rare.[4] Clinical advice to the EAG noted that in some hospitals, patient presenting with an overtly malignant feeling prostate gland (T4) and high PSA may proceed directly to TRUS and biopsy before having MRI to speed up diagnosis. Reasons for deviating from the recent NICE guidance include challenges in meeting waiting targets and the limited availability of mpMRI slots. The COVID-19 pandemic has also disrupted the implementation of the guidance.[23, 24]

Clinical advisers to the EAG highlighted that bpMRI is sometimes used in current practice where mpMRI is not available. Although the 2019 National Prostate Cancer Audit (NPCA) indicated that 98% of NHS organisations were able to offer mpMRI on site, challenges in meeting the 28-day diagnostic waiting target have been reported.[25] However, there is no evidence that the accuracy of mpMRI and bpMRI differ in treatment-naïve patients.[26]

Although uptake of mpMRI as first-line diagnostic test has increased in recent years, it is unclear to what extent this is implemented in the NHS, and whether and to what extent other alternative pathways may be followed.

2.2.3 Biopsy

The decision to collect biopsy samples is informed by the MRI, as well as specific risk factors (such as PSA density, family history and ethnicity) and individual clinician preference. One or more prostate biopsies may be performed to rule out or confirm the presence of PCa. Different methods exist for sampling the prostate tissue. The site(s) for biopsy can be *targeted* for people who have a suspicious lesion identified by the MRI scan. MRI-influenced biopsies, or targeted biopsies, are biopsies where the site (or sites) for biopsy are targeted based on the location of one or more suspicious lesions identified by the MRI scan. Tissue samples or cores are only collected from the areas identified in the MRI scan as suspicious. The biopsies can also be *systematic*, where multiple samples are taken in a systematic fashion from different regions of the prostate according to a predefined scheme rather than guided by the MRI results. A systematic only biopsy approach may be taken for instance where clinical suspicion is high but not reflected in the MRI (typically with a Likert/PI-RADS score of 2 or less), although there is regional variation in this practice.

Prostate biopsies may be performed via the transrectal route or the transperineal route. Both routes use a transrectal ultrasound probe inserted into the anus to generate a live image of the prostate. With TRUS prostate biopsy, a biopsy needle is inserted through the rectal wall via the anus. TRUS biopsies are usually performed under local anaesthesia, although it can also be carried out under general anaesthesia (for example if the patient is unlikely to tolerate the procedure otherwise). In a transperineal biopsy, the biopsy needle is inserted through the perineum. Historically, transperineal biopsies were always conducted under general anaesthesia. However, recent developments in transperineal biopsy techniques have made the procedure more tolerable, and it is now routinely performed under local anaesthesia.[27] NICE draft guidance has recently recommended local anaesthetic transperineal (LATP) prostate biopsy, using the freehand needle positioning devices PrecisionPoint, EZU-PA3U device, Trinity Perine Grid, and UA1232 puncture attachment, as options for diagnosing PCa.[28, 29] Furthermore, patients may receive a spinal block prior to the biopsy being taken, although practice will vary between centres. Spinal anaesthesia may be conducted in an outpatient office[30] or operating theatre.[31]

When a prostate biopsy is performed, tissue cores from the prostate are obtained for histological examination. The number of cores sampled primarily depends on the biopsy technique, but may also vary based on whether the patient has a previous negative biopsy. In a systematic biopsy, the number of cores sampled can range from six, to 12 or 14. When more samples are obtained, a greater volume of the prostate gland is sampled, potentially increasing the detection rate. Obtaining any further cores is associated with a limited increase in diagnostic yield,[32] but an increased risk in the incidence of complications, such as bleeding (haematuria, haematospermia, haemoejaculate, haematochezia or rectal bleeding), infections (e.g. urinary tract infection), pain, urinary retention and erectile

dysfunction.[33] In MRI guided biopsies, fewer cores can be obtained, as sampling can be targeted at the areas where there is a high-suspicion of cancer. The NICE guidelines do not specify the number of cores that should be obtained from each suspicious area; European guidelines state that multiple (three to five) biopsy cores per lesions should be taken to reduce the chance of missing or under sampling lesions,[34] whereas guidance from the American Urological Association and the Society of Abdominal Radiology's Prostate Cancer Disease-Focused Panel[35] notes that at least two target cores per region of interest should be obtained. Clinical advisers to the EAG indicated that a minimum of two cores per targeted lesion were typically taken in NHS practice, and that for most patients, only one lesion (typically the largest) was targeted.

NICE NG131 recommend that a targeted, MRI-influenced prostate biopsy should be offered to people whose Likert score is 3 or more.[10] Currently, MRI-influenced prostate biopsy may use one of three different approaches:

- Cognitive fusion (or visual estimation), in which the operator interprets the MRI imaging before the biopsy and manually targets the area of interest using TRUS as a guide; additional samples are also taken in a systematic way according to a pre-defined protocol.
- Software fusion, which automatically overlays the MRI image onto the real-time TRUS therefore allowing for real-time visualisation of the area of interest where targeted samples are taken; additional samples are also taken in a systematic way according to a pre-defined protocol.
- In-bore biopsy, or 'in-gantry' biopsy, a technique that involves performing the prostate biopsy in the MRI scanner, where the diagnostic MRI is fused with real-time MRI using the MR images taken immediately after each needle placement to guide the biopsy.

Cognitive fusion is the current standard of care. Clinical advisers to the EAG noted that different versions of software fusion are currently used in a number of NHS centres. In-bore biopsies, and MRI fusion software that integrates AI-driven diagnosis of PCa, are not used in standard clinical practice. MRI-influenced prostate biopsy methods are further discussed in sections 1.4 and 1.5.

Software fusion and cognitive fusion prostate biopsy can be performed with or without the addition of systematic biopsy. The European Association of Urology (EAU) guidelines on PCa recommends combining targeted and systematic biopsy in people with a PI-RADS score of 3 or more who have not had a prior biopsy.[34] In UK clinical practice, after targeting sites of interest for biopsy in eligible people, additional biopsy cores may be taken from the area around the target lesion and a systematic biopsy is performed in addition to the targeted biopsy. Although not strictly recommended by NICE, their guideline on the diagnostic and management of PCa (NG131) notes that most often, MRI influenced biopsies will be performed in combination with systematic biopsies.[10] However, there is

variation in practice dependent on local protocols in terms of whether off-target cores are sampled or not, the number of samples taken and the sampling pattern for the systemic component of combined biopsies. For people whose Likert score is 1 or 2, omitting a prostate biopsy should be considered but only after discussing the risks and benefits with the person and reaching a shared decision. If a person opts to have a biopsy, systematic prostate biopsy (whereby multiple samples are taken in a systematic fashion from different regions of the prostate according to a predefined scheme) is offered. NHS England guidance[17] states that people with a Likert or PI-RADS score of 1 or 2 and people with a Likert or PI-RADS score of 3 who also have a PSA density less than 0.15ng (or 0.12ng in some centres) of PSA per mL of serum per mL of prostate volume may be discharged, taking account of risk factors and patient preferences.

For those patients whose MRI influenced biopsy is negative, results will be reviewed by a urological cancer multi-disciplinary team (MDT) typically including a urologist and a radiologist, and the possibility of significant disease discussed with the patient. However, clinical advice to the EAG noted that in practice, not all hospitals are able to perform an MDT review of all negative MRI influenced biopsies, in which case results may be sent for individual clinician review. A decision to offer a repeat biopsy is based on individual risk factors, including whether the biopsy showed high-grade prostatic intra-epithelial neoplasia, atypical small acinar proliferation or whether the digital rectal examination result was abnormal.[12, 17, 34] Clinical advice to the EAG noted that factors determining eligibility for, and timing of, repeat biopsy may vary across centres and will depend on individual risk factors, although patients with a negative biopsy and PIRADS/Likert scores of 4 or 5, larger suspicious lesions on MRI and fitter patients are more likely to undergo repeat biopsy within 12 months. If a repeat biopsy is not offered, patients could instead undergo active surveillance with PSA testing or may be discharged depending on MRI and histology findings.[17] Patients whose repeat biopsy result is positive may be offered active surveillance or radical treatment, depending on individual patient characteristics and preferences (see Section 2.2.3.1). Patients with a negative repeat biopsy may be discharged, or have their PSA levels monitored if cancer is still suspected. Antibiotics combined with PSA monitoring may be administered to rule out prostatitis, which may show as false positive on MRI. In some rare cases, further tests such as an additional repeat biopsy, template biopsy, or a positron emission tomography (PET) scan may be conducted to definitely rule out cancer.

Following the biopsy, a pathologist will look at the biopsy samples and assign a Gleason score. The Gleason score is a grading system which estimates the aggressiveness of the PCa, based on the pattern of the cancer cells and the extent of cell differentiation. Gleason grade 1 cells look like normal prostate tissue, and Gleason grade 5 cells have mutated to such an extent that they do not resemble typical prostate cells. A primary grade is given to describe the cells that make up the largest area of the tumour and a secondary grade is given to describe the cells of the next largest area. For example, a

Gleason Score written as 3+4=7 indicates that most of the tumour is grade 3 and the next largest section of the tumour is grade 4. The two most common patterns of cells (for example, Gleason grade 3 and 4) are added together to determine a Gleason score. Gleason scores can range from 2-10, with a score of 6 being the lowest grade cancer. To help with outcome prediction and patient communication, Gleason scores ≤6, 3+4, 4+3, 8 and 9-10, respectively, can be reported as five risk groups defined by the International Society of Urological Pathology (ISUP), i.e. ISUP grades 1-5, respectively.[36]

Although the exact definition of clinically significant PCa varies across studies, it commonly refers to organ-confined cancer above a specific Gleason score (or grade) and maximum cancer core length, indicating PCa that may cause excess morbidity or death.[34] European guidelines state that lesions with a Gleason score between two and six can be considered clinically insignificant. Recent studies have commonly defined clinically significant PCa as above a Gleason score of 7 (3+4), some have used a narrower definition, including above 7 (4+3).[19, 37-39] Some publications provide more than one definition within a single study, reflecting the lack of consensus and difficulty in defining clinical significance. [40, 41]

People diagnosed with PCa are assigned a Cambridge Prognostic Group (CPG) risk category. The CPG score is assigned based on the person's PSA levels, the Gleason score of the lesion(s) (based on histological analysis of the biopsy) and the clinical stage of the disease.[10] The EAU guidance states that further tests, such as abdominopelvic imaging and bone scans, may be required to determine clinical stage of the disease when there is suspicion that the cancer has spread to the lymph nodes or the bone marrow.[34] The CPG risk category and definition is described in Table 1.

Table 1. Cambridge Prognostic Group risk categories and the respective definition based on Gleason Score, PSA level and clinical stage.

Cambridge Prognostic Group (CPG)	Risk Category	Definition	Management Recommendations (NICE NG131)
CPG 1	Low risk	Gleason score 6 (GG 1) *AND* PSA < 10 ng/ml *AND* stages T1-T2	• Offer active surveillance • Consider radical prostatectomy or radical radiotherapy if active surveillance is unsuitable or not acceptable to the person.
CPG 2	Favourable intermediate risk	Gleason score 3+4 = 7 (GG 2) **OR** PSA 10-20 ng/ml *AND* stages T1-T2	• Offer a choice between active surveillance, radical prostatectomy, or radical radiotherapy if radical treatment is suitable
CPG 3	Unfavourable intermediate risk	Gleason score 3+4 = 7 (GG 2) *AND* PSA 10-20ng/ml *AND* stages T1-T2 **OR** Gleason score 4 + 3 = 7 (GG3) *AND* stages T1-2	• Offer radical prostatectomy or radical radiotherapy • Consider active surveillance for people who choose not to have immediate radical treatment.
CPG 4	High risk	**One of:** Gleason score 8 (GG 4)	

		OR PSA >20 ng/ml **OR** Stage T3	For individuals with localised and locally advanced prostate cancer: • Do not offer active surveillance • Offer radical prostatectomy or radical radiotherapy
CPG 5	Very high risk	**Any combination of:** Gleason score 8 (GG 4), PSA > 20ng/ml or Stage T3. **OR** Gleason score 9-10 (GG 5) **OR** Stage T4	

CPG, Cambridge Prognostic Group; GG, Gleason Grade; PSA, prostate specific antigen

These risk categories, along with the outcome of discussion with patients regarding the benefits and harms of the treatment options, determine which treatment option is chosen. This ranges from active surveillance, for patients with CPG 1 or 2, to radical prostatectomy or radical radiotherapy for people with localised cancer and CPG ≥2. Patients with locally advanced PCa and CPG 4 or 5 may also be offered docetaxel chemotherapy. The recommendation to use the CPG five-tier risk prediction model was included in the NICE NG131 2021 update[10] and superseded a three-tier risk classification model including low-, intermediate- and high-risk/locally advanced groups, which did not differentiate between favourable intermediate risk (CPG 2) and unfavourable intermediate risk (CPG 3). Another important difference between the two classifications is that CPG 1 includes more men than the low-risk group in the previously recommended risk classification; some men who previously would have been in the intermediate-risk group are now classified as CPG 1. This change in risk prediction model aims to reduce under- and over-treatment in people who are at either end of the tiers, following evidence from the NICE's surveillance programme that indicated that active surveillance may not be appropriate in patients with unfavourable intermediate PCa, and that patients with favourable intermediate risk and lower risk may be over-treated.[10, 12, 42, 43]

2.2.3.1 Software fusion prostate biopsy

Using a digital overlay, software fusion biopsies allow operators to view a real-time ultrasound image alongside the patient's MRI. This requires a period of preparation, to obtain and annotate the MRI images prior to biopsy.[44] MRI images are first downloaded onto a dedicated processing software before they are annotated by contouring the edge of the prostate and the regions of interest. Clinical advice to the EAG suggests that, for an experienced practitioner, this contouring can take around 5-7 minutes. The annotated MRI scans are then uploaded onto a fusion software platform and are fused with the real-time ultrasound image. Updates to the fusion software are possible, and depending on the fusion device, are covered by a service contract or can be purchased with a one-off payment.

Use of software fusion prostate biopsy systems may potentially improve detection rate of clinically significant PCa compared with cognitive fusion, whilst reducing the number of samples taken, potentially reducing pain and risk of sepsis associated with the procedure. It could improve the

accuracy of assignment of prognostic scores such as Gleason, which influences subsequent treatment and associated patient outcomes. The technology could reduce the number of repeat biopsies for those patients with a negative index biopsy, avoiding unnecessary travel and anxiety for the person. Some fusion technologies also allow operators to keep records of previous biopsy sites to allow the urologist to return to those areas with greater precision for follow-up or additional testing.

However, the accuracy of a prostate biopsy may be impacted by a number of factors. Movement during the procedure (which could stem from patient pain),[45] operator experience,[46] difference in bladder size or prostate deformation may impact the accuracy of the biopsy, as the MRI image may not accurately reflect the prostate shape at the point of biopsy. Mechanisms using 'elastic' prostate registration, where the MRI image alters to fit the ultrasound image, have been designed to account for prostate deformation and allow for more accurate targeting of the lesions of interest.[47] Errors during the fusion of images, specifically incorrect image registration or discordance between the MRI and ultrasound image planes, especially around the base of the prostate, can lead to biopsy failure.[48]

The mechanism by which software fusion techniques may lead to improved accuracy relates notably to a better targeting of suspicious prostate lesions, including in locations that are more challenging to diagnose, such as anterior and posterior lesions.[49, 50] However, evidence for the accuracy of software fusion biopsy systems compared with cognitive fusion methods is limited. Watts (2020)[51] and Sathianathen (2019)[52] found no statistically significant difference between software fusion and cognitive fusion in prostate cancer detection, whilst Bass (2021)[53] found no evidence that software fusion was superior to cognitive fusion at detecting clinically significant prostate cancers. An older review found that software fusion biopsies detect more clinically significant cancers, using fewer biopsy cores.[54] Between-study heterogeneity ranged from moderate[51] to high,[53] although review methods and selection criteria varied.

2.2.4 Prostate cancer management: active surveillance, watchful waiting and radical treatment options

Active surveillance is a monitoring strategy for people with localised PCa for whom radical treatments (such as radical prostatectomy or radical radiotherapy) are suitable; it allows avoiding or deferring these treatments when disease progression is likely to be slow, whilst maintaining the possibility to initiate timely curative treatment. Current NICE guidance suggests a schedule of active surveillance involving regular monitoring of PSA levels and kinetics, and annual digital rectal examinations. Reassessment with mpMRI and/or re-biopsy can be triggered if concerns about clinical or PSA changes emerge at any time during active surveillance; a positive result (Gleason score 3+4 or above) on re-biopsy would then result in offering radical treatment.

For people with CPG 1, active surveillance is offered (radical treatments can be considered if active surveillance is not suitable or acceptable to the person). For people with CPG 2, a choice between radical radiotherapy with androgen deprivation (anti-hormone therapy), radical prostatectomy or active surveillance is given. For people with CPG 3, localised PCa, radical prostatectomy or radical radiotherapy with androgen deprivation is offered, and active surveillance can be considered for people who choose not to have immediate radical treatment. This recommendation is informed by a randomised trial that found that prostate cancer-specific mortality is low (approximately 1%) at 10 years follow-up and does not differ significantly between active surveillance, prostatectomy or radical radiotherapy in individuals with localised PCa, although surgery and radiotherapy resulted in lower incidences of disease progression and metastatic disease compared with active monitoring. Radical prostatectomy may also be associated with worse urinary and erectile dysfunction outcomes compared with active surveillance and radical radiotherapy at up to six years follow-up.[55] People with CPG 4 and 5 localised or locally advanced PCa should be offered a combination of radical radiotherapy and androgen deprivation. Evidence from an individual patient data (IPD) meta-analysis shows that the addition of androgen deprivation therapy to radiotherapy significantly improves metastasis-free survival.[56] Brachytherapy (a form of radiotherapy where radiation is directly targeted on the tumour by inserting radioactive pellets into the prostate) in combination with external beam radiotherapy should also be considered for people with CPG 2, 3, 4 and 5 localised or locally advanced PCa.[57] Randomised controlled trial (RCT) evidence shows a reduction in biochemical failure (such as local recurrence or distant metastases) associated with the use of low dose rate brachytherapy plus external beam radiotherapy at 6.5 years follow-up for people with high-risk (CPG 4 and 5) localised PCa.[58]

Radical prostatectomy or radical radiotherapy is offered to people with CPG 4 and 5 localised and locally advanced PCa, when it is likely the person's cancer can be controlled in the long term. Docetaxel chemotherapy may also be considered for these patients. This recommendation follows RCT evidence indicating that clinical progression-free survival (PFS) was prolonged in individuals with hormone-sensitive high risk PCa receiving docetaxel compared to standard care alone.[59-61]

Finally, some patients with metastatic disease, where the cancer has spread outside the prostate may still undergo targeted biopsy to aid decision making for localised treatment where the patient may receive some symptomatic benefit.

People with localised PCa who do not wish to undergo potentially curative treatment with radical prostatectomy or radical radiotherapy (or for whom this is not suitable), can be managed with watchful waiting. This is a monitoring strategy that aims to achieve disease control rather than cure. It is less intensive than active surveillance and involves fewer tests (e.g., annual PSA level measurements not leading to MRI or biopsy[10]) and is typically offered to older, frailer populations. Since MRI as first-line test is only recommended for patients fit for radical treatment, only a small

subset of patients who received an MRI for suspected prostate lesions, such as those with worsening health since initial investigation and a PCa diagnosis, are expected to undergo watchful waiting in practice. Some patients who are not fit enough or eligible for curative treatment may also be offered an MRI because their lack of eligibility for radical treatment is not identified prior to undergoing imaging.

2.3 Description of technologies under assessment

This assessment will evaluate software fusion technologies matching the following criteria:

- intended for use in people with suspected prostate cancer;
- available in the UK;
- holds a CE-mark;
- compatible with MRI scanners of 1.5 Tesla field strength or above;

This includes; Artemis (InnoMedicus Artemis), Biojet (Healthcare Supply Solutions Ltd), BiopSee (Medcom), bkFusion (BK Medical UK Ltd and MIM Software Inc), Fusion Bx 2.0 (Focal Healthcare), FusionVu (ExactImaging), iSR'obot^{TM} Mona Lisa (Biobot iSR'obot), KOELIS Trinity (KOELIS and Kebomed), and UroNav Fusion Biopsy System (Phillips). Table 2 presents a brief summary of the characteristics of these nine technologies.

Software fusion devices can have a variety of different features, which means they vary in the way in which they operate.

- **Positioning of the ultrasound probe:** An ultrasound probe can be cradled and held stationary using a device called a stepper which is attached to a workstation (also known as a stabilised approach). It can be supported by a semi-robotic arm, which allows for the ultrasound probe to be manoeuvred, whilst maintaining a stable pressure on the prostate. The semi-robotic arm can be used as a stepper for stabilised biopsies or can allow complete freedom of movement for use during a freehand biopsy. Finally, the ultrasound probe can be held by hand (using a freehand technique).
- **Core sampling technique:** Different techniques can be used to take the cores, especially in the case of transperineal biopsies. First, a grid or template can be used, which is attached to a stepper and placed in front of the perineum. The grid is marked with a number of holes, which correspond to a letter and a number to allow for multiple cores to be taken in a systematic way. Alternatively, a coaxial needle can be used. In this technique, a larger introductory needle is used to puncture the perineum before the biopsy needles is passed through. This biopsy needle can be angled to take multiple biopsies without creating multiple puncture wounds to the perineum. The coaxial needle is used with the freehand technique, where it is

attached to the ultrasound probe, or in a double freehand technique, where the needle is held by hand.

- **Image registration:** During software fusion, the mpMRI images are fused with the ultrasound images during the biopsy procedure. The mpMRI image can be fixed (known as rigid registration) and will not move when the prostate is deformed, either by patient movement, or by the insertion of a needle; or elastic, which means the mpMRI image adjusts to match the ultrasound image to account for prostate deformation.

A description of the principal features of the technologies is given below.

2.3.1 Artemis (InnoMedicus Artemis)

The Artemis fusion biopsy system comprises a semi-robotic mechanical arm and a mobile workstation. The system includes the ProFuse radiology software for preparation of MRI data for fusion and reporting findings on the ARTEMIS biopsy system. The system allows both elastic and rigid estimation to account for prostate deformation, and supports both transrectal and transperineal biopsies. The mechanical arm is used to track the prostate in real-time and guide the biopsy needle.

At the time of writing the EAG report, the company had not registered with NICE, and therefore did not provide information on this technology's compatibility with a picture archiving and communication system (PACS), image measurement capabilities and ability to produce archivable cartograms.

2.3.2 Biojet (Healthcare Supply Solutions Ltd)

The BioJet MR Fusion system comprises MRI fusion software, a mobile workstation, and is compatible with third party ultrasounds. The system uses elastic estimations and is compatible with both transrectal and transperineal biopsies and supports both stabilised and freehand biopsy approaches.

The software enables image measurements and generates reports displaying the location of sampled areas. BioJet can be connected to a local PACS.

2.3.3 BiopSee (Medcom)

The BiopSee consists of the BiopSee software and the MedSta cart (workstation), and is compatible with third party ultrasounds. The system supports both elastic and rigid estimation to account for prostate deformation, and allows both transrectal and transperineal biopsies. The system can be used for stabilised and freehand biopsy approaches. A stabilising arm is available for transperineal stabilised biopsies. Patient movement is tracked through the stepper during stabilised biopsies, or

through a magnetic tracker, which is attached to the probe during freehand biopsies. The system can automatically adjust for patient movement, or the user can manually adjust the contours when a patient moves.

The BiopSee records all positions of the needle and shows the coverage of the prostate. Image measurements such as prostate and lesion volumes are also possible. The data is stored locally and can be connected to a PACS for import and export of images.

2.3.4 bkFusion (BK Medical UK Ltd and MIM Software Inc)

BK Medical UK Ltd offers three versions of bkFusion software: one for transrectal, one for freehand transperineal and one for stabilised transperineal biopsies. The software can be integrated into either the bk3000 or bk5000 ultrasounds. The bkFusion system uses rigid estimation to account for prostate deformation. Predictive Fusion software re-orientates the MRI image before the biopsy. The transrectal and freehand transperineal fusion systems comprises a magnetic field generator and sensor to track the probe position.

Image measurements such as prostate volume are possible. A detailed report of the biopsy can be saved locally, or transferred to a PACS.

2.3.5 Fusion Bx 2.0 (Focal Healthcare)

The Fusion Bx 2.0 is a biopsy device that includes a counter-balanced, semi-robotic arm that is mounted to a mobile cart. The Fusion Bx 2.0 comprises Fusion MR software which is compatible with third party ultrasounds. The system uses both elastic and rigid estimation to account for prostate deformation, and supports both transrectal and transperineal biopsies. Patient movements are tracked with sensors inside the semi-robotic arm.

The software allows image measurements such as prostate volume and distances can be calculated. Data on the biopsied samples and the regions of interest are recorded on a 3D image of the prostate. The system can connect to PACS using a wired Ethernet or Wi-Fi connection.

2.3.6 FusionVu (ExactImaging)

The ExactVu micro-ultrasound device includes a FusionVu feature that enables software fusion biopsy. A stabiliser arm or stepper is available for stabilised biopsies, and freehand biopsies are also possible. The system uses rigid estimation followed by real-time visualisation of the lesions using micro-ultrasound, and supports both transperineal and transrectal biopsies. The system tracks and adjusts for patient movement using data from a movement sensor together with the live ultrasound images.

The software provides image measurements such as prostate volume and lesion size. Information on the orientation of all images and video frames are recorded so that the same position can be found if a repeat biopsy is performed. The system is PACS compatible, but a separate software (Weasis DICOM viewer) is available where a PACS is not available.

2.3.7 iSR'obot Mona Lisa (Biobot iSR'obot)

The iSR'obot Mona Lisa is a robotic transperineal prostate biopsy system with MRI-ultrasound fusion capability. The system uses UroFusion software to highlight regions of interest on MR images and fuses the MRI model with the ultrasound model. The robotic needle guide allows automated positioning and depth control of the biopsy needle to the targeted biopsy core. The system uses elastic estimation to account for prostate deformation.

Reports are generated with 3D images and coordinates are recorded of each biopsy sample. At the time of writing the EAG report, the company had not registered with NICE, and therefore did not provide information on the tracking of patient movement, whether freehand biopsies can be done, PACS compatibility and image measurement capabilities of this system.

2.3.8 KOELIS Trinity (KOELIS and Kebomed)

The KOELIS Trinity is a mobile ultrasound system with mapping fusion software, which comprises PROMAP 3D-Prostate Suite software and the TRINITY ultrasound system (workstation, RECFIRE ultrasound probes, guides specific to transperineal or transrectal biopsies, and a Steady Pro probe holder). The system uses elastic estimation to account for prostate deformation, and supports both transrectal and transperineal biopsies. It enables both stabilised and freehand probe biopsies. The Organ Based Tracking Fusion software identifies and compensates for patient movements and prostate deformations to record each core location.

The PROMAP software produces a 3D map of the prostate recording the position of MRI lesion targets and location of biopsy samples. The KOELIS Trinity provides image measurements such as prostate volume, exact measurements of the regions of interest and other quantitative measurements of the image. Data can be transferred to a PACS.

2.3.9 UroNav Fusion Biopsy System (Phillips)

The UroNav Fusion Biopsy System includes an electromagnetic tracking system, a mobile workstation and DynaCAD Prostate fusion software. The system is compatible with third party ultrasounds. It supports both transperineal and transrectal biopsies, with stabilised or freehand approaches. The system can be used with the UroNav mobile stepper system and the two navigation sensors to track patient movement.

The UroNav Fusion Biopsy system provides the core location data, images and videos. At the time of writing the EAG report, the company had not registered with NICE, and therefore did not submit any information on image estimation methods for prostate deformation, patient movement tracking feasibility for freehand biopsies, PACS compatibility and image measurement capabilities of this system.

Table 2. Summary of technologies features

Software system	Manufacturer	Hardware system	Fixation for biopsies	Elastic or rigid estimation	Was there a submission for the DAR?
ARTEMIS	InnoMedicus Artemis	ARTEMIS	Stabilised, Freehand unknown, Semi-robotic arm	Both	No
BioJet	Healthcare Supply Solutions Ltd	Third party ultrasounds	Stabilised, Freehand (without tracking movement)	Both	No
BiopSee	Medcom	MedSta or third-party ultrasounds	Stabilised, Freehand	Both	Yes
bkFusion	BK Medical UK Ltd and MIM Software Inc	BK3000 or BK5000	Stabilised, Freehand	Rigid	Yes
Fusion Bx 2.0	Focal Healthcare	Third party ultrasounds	Stabilised, Freehand, Robotic arm	Both	Yes
Fusion Vu	ExactImaging	ExactVu	Stabilised, Freehand	Rigid	Yes
iSR'obot™ MonaLisa	Biobot iSR'obot	iSR'obot™ Mona Lisa	Stabilised, Freehand unknown, Robotic arm	Elastic	No
KOELIS Trinity	KOELIS and Kebomed	TRINITY ultrasound system	Stabilised, Freehand	Elastic	Yes
UroNav Fusion Biopsy System	Phillips	Third party ultrasounds	Stabilised, Freehand	Unknown	No

2.3.10 Other interventions

'In-bore' biopsy, or 'in-gantry' biopsy, is a technique that involves performing the prostate biopsy in the MRI scanner, using the MR images taken immediately after each needle placement to guide the biopsy. The use of in bore MRI and AI-driven software are beyond the scope of this assessment.

2.4 Place of the intervention in the diagnostic and care pathway

Software fusion targeted biopsy for people with suspected PCa takes place at the same two points in the diagnostic pathway as targeted cognitive fusion biopsy, the current standard of care.

Patients having a first targeted biopsy

Software fusion biopsy (with or without systematic biopsy) would be offered as an alternative to targeted cognitive fusion biopsy to people with a Likert/PI-RADS score of 3 or more following an MRI, after having been referred to secondary care with suspected PCa (with PSA levels above the age-specific reference range or those whose prostate is suspicious of malignancy based on rectal examination). Clinical advisers to the EAG indicated that biopsy-naïve patients represented the large majority (more than 90%) of patients with suspected PCa undergoing targeted biopsy.

Patients having a repeat targeted biopsy

Patients offered a repeat biopsy following a prior negative biopsy could also be offered software fusion biopsy as an alternative to targeted cognitive fusion. As discussed in in Section 2.2, NG131 recommends that an MDT decides on whether to offer a repeat biopsy based on individual risk factors, although not all centres may be able to perform an MDT review of all negative MRI-influence biopsies, and eligibility and timing of repeat biopsy may vary in practice. In clinical practice, repeat biopsies are likely to be offered to patients whose mpMRI results were not consistent with the biopsy (i.e. mpMRI of 4-5 and no PCa detected on biopsy). NG131 does not recommend repeat MRI for patients requiring a repeat biopsy, instead a repeat targeted biopsy can be conducted based on the initial MRI report. EAG clinical advisers suggested this subgroup would make up less than 10% of patients with suspected PCa.

Potential pathway positions out of scope for the current assessment

Although software fusion may also be used to monitor patients and inform treatment for individuals with a PCa diagnosis in active surveillance, this population is beyond the scope of this assessment.

2.5 Relevant comparator

The comparator for this assessment is targeted transperineal or transrectal prostate biopsy using cognitive fusion with or without systematic biopsy, under local or general anaesthesia, in which the operator interprets the MRI imaging before the biopsy and manually targets the area of interest using TRUS as a guide. Clinical advisers to the EAG highlighted that the expertise of the person performing the biopsy may affect the accuracy and procedure time of cognitive fusion.

3 AIMS AND OBJECTIVES

The aim of the project was to assess the clinical and cost-effectiveness of software fusion biopsy systems in people with suspected localised and locally advanced PCa, by addressing the following protocol-specified objectives:

Clinical effectiveness

- To perform a systematic review of the diagnostic accuracy and clinical efficacy of nine software fusion systems compared with cognitive fusion targeted biopsy and with each other, in people with suspected PCa who have had an MRI scan that indicates a lesion.

- To compare the diagnostic accuracy of different software fusion biopsy systems with each other and with cognitive fusion targeted biopsy in people with suspected PCa who have had an MRI scan that indicates a lesion using meta-analytical methods and to combine the diagnostic accuracy of different software fusion systems where appropriate.

- To perform a narrative systematic review of the clinical efficacy, safety and practical implementation of software fusion targeted biopsy. This includes assessment of intermediate outcomes, mortality and morbidity, patient-centred outcomes, adverse events, and acceptability to clinicians and patients.

Cost effectiveness

- To conduct a systematic review and critical appraisal of relevant cost-effectiveness evidence of the use of software fusion biopsy systems compared to cognitive fusion for targeted biopsy in people with suspected PCa who have had an MRI scan indicating a lesion.

- To develop and validate a decision-analytic model to estimate the cost-effectiveness of software fusion targeted biopsy systems in people with suspected PCa who have had an MRI scan indicating a lesion compared to targeted biopsy using cognitive fusion. This will require linking intermediate outcomes, such as the diagnostic accuracy of software fusion biopsy systems to subsequent management decisions and to final health outcomes including morbidity and mortality associated with alternative treatment options (e.g., active surveillance and radical treatment). Final health outcomes will be evaluated in terms of quality-adjusted life years (QALYs).

- To populate the model using the most appropriate available evidence. This evidence is likely to be identified from published literature, routine data sources and potentially using data elicited from relevant clinical experts and companies.

- To estimate the incremental cost-effectiveness of the software fusion biopsy systems compared to the current standard of care for the population of interest (cognitive fusion biopsy), based on an assessment of long-term NHS and Personal Social Service (PSS) costs and quality-adjusted survival. The time horizon of the model will be sufficient to capture both the short-term and longer-term outcomes.

- To characterise the parameter uncertainty in the data used to populate the model and present the resulting uncertainty in the results to decision makers. To this purpose, we will perform comprehensive (probabilistic and deterministic) sensitivity analyses varying parameter inputs, and structural assumptions of the model, as appropriate.

- Where possible and applicable, to assess the impact of potential sources of heterogeneity on cost-effectiveness, including subgroup analyses (e.g., patients with previous negative biopsy results within 12 months) and consideration of other factors that may affect diagnostic accuracy.

4 ASSESSMENT OF DIAGNOSTIC ACCURACY AND CLINICAL EFFECTIVENESS

This section presents the methods and results of the systematic review of diagnostic accuracy and clinical effectiveness. Section 4.1 details the systematic review methods, and Section 4.2 presents the data synthesis methods. Section 4.3 summarises the quantity and quality of evidence included in the systematic review, Section 4.4 presents the diagnostic accuracy results of the systematic review and meta-analysis; results for all other outcomes included in the systematic review are presented in Section 4.5. Section 4.6 summarises the key findings from the systematic review, and Section 4.7 presents a summary of additional evidence identified to inform the economic model.

4.1 Systematic review methods (study selection, data extraction, quality assessment)

4.1.1 Searches

The aim of the literature search was to systematically identify published and unpublished studies of prostate biopsies utilising either software fusion or cognitive fusion.

An information specialist (MH) developed a search strategy in Ovid MEDLINE using textword searches of the title and abstracts of database records along with relevant subject heading searches. The search strategy consisted of: 1) terms for prostate cancer AND 2) terms for MRI AND 3) terms relating to fusion techniques AND 4) terms for prostate biopsy. The terms used to describe fusion techniques was found to vary in the literature with some articles lacking any terms for fusion techniques in the title, abstract or subject headings of the database record. Therefore, related terms such as targeted biopsy, focal biopsy or MRI guided biopsy were added to the strategy along with some proximity searching to capture phrases in the title and abstracts of records around the use of MRI prior to a prostate biopsy. Named software fusion software and hardware was also included in the strategy (e.g., Fusion Bx, Biojet, KOELIS Trinity, bkFusion).

A date limit was applied (from 2008 onwards), due to the relatively recent nature of the technologies under assessment, and as informed by results of scoping searches and previous systematic reviews.[51, 53, 62, 63] No language or study design restrictions were applied to the searches. The MEDLINE strategy was agreed with the review team and checked by a second information specialist using aspects of the PRESS checklist.[64] The final MEDLINE strategy was adapted for use in all resources searched.

The following databases were searched in May 2022: MEDLINE ALL (Ovid), Cochrane Controlled Register of Trials (Wiley), Cochrane Database of Systematic Reviews (Wiley), Cumulative Index to Nursing & Allied Health (Ebsco), Database of Abstracts of Reviews of Effects (CRD databases), EconLit (Ovid), Embase (Ovid), Health Technology Assessment database (CRD databases), Health Management Information Consortium (Ovid), International Health Technology Assessment (INAHTA) database, Latin American and Caribbean Health Sciences Literature (LILACS) database, NHS Economic Evaluation Database (CRD databases), Science Citation Index (Web of Science).

Further ongoing and unpublished studies were identified through searches of: ClinicalTrials.gov, Conference Proceedings Citation Index: Science (Web of Science), European Union Clinical Trials Register, Open Access Theses and Dissertations, Proquest Dissertations & Theses A&I, PROSPERO, WHO International Clinical Trials Registry Platform portal.

A search for relevant guidelines was carried out via the following websites: National Institute for Health and Clinical Excellence (NICE), ECRI Guidelines Trust, Guidelines International Network (GIN) international guideline library and the Trip database. Full search strategies for all resources can be found in Appendix 1.

Additionally, company websites were searched to identify relevant publications and other materials relating to the technology, and companies registered with NICE at the time of the protocol submission were contacted for further details about their respective technologies. Reference lists of included studies and relevant systematic reviews were scanned to identify any further potentially relevant studies.

An update search was carried out on 2nd August 2022 to capture any recently published studies. The update search was undertaken on the following four databases: MEDLINE ALL (Ovid), Cochrane Controlled Register of Trials (Wiley), Embase (Ovid) and the Science Citation Index (Web of Science). Search results were downloaded from each database and added to the EndNote library of original search results for deduplication.

4.1.2 Selection criteria

All titles and abstracts were screened independently by two reviewers (AL and LB). Full text papers of any titles and abstracts deemed to be relevant were obtained where possible, and the relevance of

each study assessed independently by two reviewers according to the criteria below. Disagreements were resolved by consensus, or where necessary, by consulting a third reviewer. Conference abstracts were considered to be eligible if they provide sufficient information for inclusion, and attempts were made to contact authors for further data. The eligibility criteria that were used to identify relevant studies are listed below.

4.1.2.1 Population

People with suspected PCa who have had an MRI scan that indicates a significant lesion (Likert or PI-RADS score of 3 or more). This included people who were biopsy naïve and those who are referred for a repeat biopsy following a previous negative prostate biopsy. No time limit since the first negative biopsy was set for inclusion of studies including patients with repeat biopsies, although applicability with respect to the scope was considered as part of the quality assessment.

Studies primarily focused on people who do not have a lesion visible on their magnetic resonance image, people on an active surveillance care pathway, and people with relapsing PCa were excluded. Patients who could not have an MRI scan were also excluded. Studies including a small subset of individuals with a Likert or PI-RADS score of 2 or less were included if they provided data primarily for the eligible population; their applicability was assessed during quality assessment.

4.1.2.2 Interventions

Studies evaluating software fusion alone or in combination with cognitive fusion or systematic biopsy, under local or general anaesthesia were eligible. No exclusions were made based on biopsy route. The following software fusion technologies were included.

- Artemis (InnoMedicus Artemis)
- Biojet (Healthcare Supply Solutions Ltd)
- BiopSee (Medcom)
- bkFusion (BK Medical UK Ltd and MIM Software Inc)
- Fusion Bx 2.0 (Focal Healthcare)
- FusionVu (Exact Imaging)
- iSR’obot Mona LisaTM (Biobot iSR’obot)
- KOELIS Trinity (KOELIS and Kebomed)
- UroNav Fusion Biopsy System (Phillips)

Where applicable, earlier versions of these technologies were also included, and their applicability was accounted for during quality assessment.

4.1.2.3 Comparators

Targeted transperineal or transrectal prostate biopsy using cognitive fusion with or without systematic biopsy, under local or general anaesthesia. Although systematic biopsies and 'in bore' biopsies are outside the scope of this review, studies that evaluate these methods were included if they provide separate data to compare targeted biopsies using software fusion against cognitive fusion. Studies evaluating several software fusion technologies against one another were also eligible for inclusion.

4.1.2.4 Reference standard

Total cancer cases in diagnostic accuracy studies are commonly identified using a combination of software fusion, cognitive fusion and systematic biopsies as 'reference standard'.[51, 53]

In those studies, diagnostic accuracy estimates of software fusion and cognitive fusion are therefore inherently dependent on the accuracy of mpMRI, TRUS and fusion approaches, as well as the accuracy of the biopsy method, which may vary by type and route. Reference standards that use histopathology from biopsy samples rather than whole prostate resection may also miss positive cases, and reference standards that include results from samples identified by software fusion and/or cognitive fusion are at risk of incorporation bias (when results of an index test are used to establish the final diagnosis). Reference standards that use histopathology from whole prostate resection are usually only reported for those who have been classified as high risk and have had radical prostatectomy. In addition, histopathology, although commonly used as gold standard test for cancer detection and grading, may also misclassify a small proportion (approximately 2%) of negative PCa cases as positive.[65]

Template-guided biopsy, including transperineal template-guided mapping biopsy (TTMB), also called template-guided saturation biopsy (TSB), is seen as a more optimal reference standard, compared with standard 12-core systematic biopsy. TTMB is a transperineal TRUS-guided biopsy of the prostate using a 5-mm brachytherapy grid, with at least one biopsy from each hole. TSB includes 20 or more transperineal or transrectal TRUS-guided biopsies of the prostate performed to comprehensively sample the whole prostate, according to a predefined core distribution pattern. Template-guided biopsies using a uniform grid and taken at 5 mm intervals can technically only miss tumours that are smaller than the distance between the adjacent cores.[66] Although template guided biopsy is imperfect, notably due to the fact that test accuracy depends on the intensity of cores taken and core trajectory,[66] it is superior to standard systematic biopsy as a reference standard as it aims to comprehensively sample all zones of the prostate. However, template-guided biopsies are invasive and may not be used in diagnostic accuracy studies, therefore combinations of reference standards with lower diagnostic accuracy (e.g., cognitive fusion with software fusion and systematic biopsies with fewer than 20 cores) were also eligible for inclusion.

A positive biopsy was defined as histopathological confirmation of one of the target conditions within the biopsy cores.

4.1.2.5 Outcomes

The following intermediate outcomes were eligible:

- Measures of diagnostic accuracy (including sensitivity, specificity, test positive/negative rates)
- Cancer detection rates (number of patients with detected cancer by software fusion or cognitive fusion divided by the total number of patients with confirmed cancer)
- Clinically significant cancer detection rates (all definitions)
- Clinically insignificant cancer detection rates (all definitions)
- Cancer detection rates by prognostic score (such as CPG 1 to 5 or other similar classification that can be mapped into the CPG classification) and/or Gleason score
- Biopsy positivity rate (ratio of positive biopsies out of total number of biopsy samples)
- Biopsy sample suitability/quality
- Number of biopsy samples taken
- Procedure completion rates
- Software failure rate
- Time to diagnosis
- Length of hospital stay (emergency department and inpatient stay)
- Time taken for MR image preparation
- Time taken for biopsy procedure
- Number of repeat biopsies within 12 months
- Subsequent prostate cancer management (such as no treatment, active surveillance, radical prostatectomy, radical radiotherapy, and hormone therapy)

The following clinical outcomes were eligible:

- Rates of biopsy related complications and adverse events, including infection, sepsis and haematuria, urinary retention, erectile dysfunction, and bowel function
- Hospitalisation events after biopsy
- Survival
- Progression free survival
- Adverse events from treatment

Patient- and carer-reported outcomes were eligible, including:

- Health-related quality of life

- Other self-reported outcomes including tolerability, embarrassment and loss of dignity

The following implementation endpoints were eligible:

- Operator preferences
- Barriers and facilitators to implementation

The following cost outcomes were eligible:

- Costs of MRI fusion software and any proprietary hardware (including the workstation, ultrasound systems, probe holders, replacement parts, consumables such as guides, and maintenance);
- Cost of staff time (including MR image interpretation time and biopsy procedure time) and of any associated training;
- Medical costs arising from the biopsy such as anaesthetic, sedation, hospital admissions and stays;
- Costs related to using intervention (including any time analysing and storing data);
- Costs of histopathology biopsy samples analysis;
- Cost of treatment of cancer (including costs of any adverse events);
- Costs relating to follow-up;
- Costs of subsequent biopsies;
- Costs arising from watchful waiting;
- Costs arising from active surveillance.

4.1.2.6 Study designs

Prospective studies comparing software fusion against cognitive fusion biopsy that report the results of both software fusion and cognitive fusion biopsy separately were considered. Studies including within-patient comparisons (where software fusion and cognitive fusion biopsy are compared within the same patient) and between-patient comparisons (where participants receive either software fusion or cognitive fusion biopsy) were included.

Where no prospective evidence could be found to inform the diagnostic accuracy of an eligible software fusion technology, retrospective studies that met all other selection criteria were included.

No restriction by healthcare setting were made.

4.1.2.7 Indirect evidence

Where the interventions of interests did not form a connected network to allow comparison of each intervention against every other, prospective, within-patient comparisons or RCTs between software fusion and systematic biopsy, and between cognitive fusion and systematic biopsy, were also eligible

to inform indirect comparisons, provided they reported numbers or rates of patients with no cancer, all PCa and clinically significant cancers for either software fusion or cognitive fusion against systematic biopsy or template biopsy, and the combination of software or cognitive fusion with systematic biopsy or template biospy.

4.1.3 Data extraction

Information on study details (including study design, sample size), patient characteristics (e.g., age, PSA, PI-RADS/Likert score and version, reason for referral, whether first biopsy, repeat biopsy and lesion location), intervention characteristics (including software fusion technology type and version, MRI technology and magnet strength, biopsy route (transrectal or transperineal) whether the procedure used fixed/free hand; local/general anaesthetic and was based on biparametric or mpMRI, the use and number of targeted and systematic core biopsy samples, operator experience), outcomes data and definitions of outcomes were extracted by at least one reviewer (AL or LB) using a standardised data extraction form and independently checked by a second reviewer (AL or LB). Discrepancies were resolved by discussion, with involvement of a third reviewer (SD) where necessary.

Where required and appropriate, attempts were made to contact companies for additional information, including unpublished data, missing data, relevant subgroup data and more granular outcome data (e.g., matrices reporting a breakdown of detection rates by cancer prognostic score). Data from relevant studies with multiple publications were extracted and reported as a single study. The most recent or most complete publication were used in situations where the possibility of overlapping populations could not be excluded. Where not reported, rates of clinically insignificant cancers were imputed by subtracting the number of clinically significant cancers from the total number of cancers detected (as per Bass 2021).[53]

4.1.4 Critical appraisal

The quality of the diagnostic accuracy studies was assessed using the tools Quality Assessment tool of Diagnostic Accuracy Studies (QUADAS)-2 and QUADAS-C tools.[67, 68] The QUADAS-2 tool evaluates both risk of bias (associated with the population selection, index test, reference standard and patient flow) and study applicability (population selection, index test and reference standard) of individual studies to the review question. The QUADAS-C tool is designed to assess risk of bias in test comparisons undertaken in studies that evaluate two or more index tests. QUADAS-C is an extension of QUADAS-2 and includes all domains covered by QUADAS-2. Each QUADAS-C domain is informed by each QUADAS-2 judgment for each test and additional signalling questions that are specific for comparisons to produce a risk of bias judgment for the comparison. The quality assessment focused on the risk of bias and applicability of cancer detection outcomes only. Since the

review focused on the relative accuracy of two index tests, QUADAS-2 risk of bias assessments were not presented. All studies were quality assessed and checked by a second reviewer. Disagreements were resolved through discussion. Decisions with rationale for judgments were presented in tables.

4.2 Data synthesis methods

4.2.1 Meta-analysis

The meta-analyses aimed to compare four types of prostate biopsy approaches: cognitive fusion, software fusion, cognitive fusion with concomitant systematic biopsies, and software fusion with systematic biopsies. When relative effects comparing more than one intervention are of interest, a network meta-analysis (NMA) should be conducted to allow comparison of all interventions to each other.[69] NMA is an extension of pairwise (two-treatment) meta-analysis to allow comparisons across more than two treatments by producing relative effects for every pair of treatments in a connected network. Direct evidence from studies comparing two interventions directly is pooled with indirect evidence from studies that have a common comparator thus allowing consistent estimates of relative effects to be produced that account for all relevant evidence and are typically more precise. Common- (fixed-) or random-effects models can be used.[70]

Since many studies compared one or more of the four biopsy types of interest to systematic biopsy alone, this biopsy type was also included in the network of interventions in order to allow more comparisons to be made and to increase precision in the estimated relative effects.[69]

NMAs were conducted using a Bayesian framework estimated through Markov chain Monte Carlo methods. In an attempt to minimise bias, only prospective studies reporting within-patient comparisons, or RCTs reporting comparative results for two or more of the interventions of interest (software fusion, cognitive fusion, systematic biopsy or a combination of software/cognitive fusion with systematic biopsy), were included in the synthesis.

Model convergence was assessed by running two independent chains with different starting values looking at history plot and through inspection of Gelman-Rubin diagnostic plots. Due to data sparseness (few studies per comparison and not all studies reporting all outcomes) only fixed-effect models were fit to the data. Model fit was assessed by comparing the mean total residual deviance to the number of independent data points contributing to the analysis.[71]

Network plots were drawn in R[72] using the *netmeta* package.[73]

4.2.2 Multinomial synthesis model

To adequately distinguish between the different biopsy methods and software fusion devices, it is necessary not only to describe how they differ in classifying patients as having PCa or not, but also how they differ in classifying patients as having PCa at different Gleason Grades, as that determines further treatment strategies. To inform post-biopsy patient management in the economic model, data are modelled by ISUP grade, where reported.

In order to best describe the differences between biopsy methods for each diagnostic category, a multinomial logistic regression model was fitted where the odds of being categorised in each of the different categories in Table 3 compared to the reference category (no PCa) are allowed to vary by biopsy type. This model is conceptually equivalent to four binomial logistic regressions comparing category $r > 1$ with category 1 (no PCa), for each different biopsy type compared to the reference, cognitive biopsy.

Table 3 Cancer detection categories used to inform the economic model

Categories	Gleason	ISUP Grade
1	-	no cancer*
2	3+3	1
3	3+4	2
4	4+3	3
5	8-10	4-5

* Although not formally part of the ISUP grade definition, it is distinguished in the model
ISUP: International Society of Urological Pathology

The multinomial logistic regression model accounts for the ordered nature of the categories, which is important since a higher or lower detection of higher-grade cancers may have an impact on the cost-effectiveness of each device. However, the model does not take into account that some of the included studies reported results from different biopsies techniques performed on the same patients.[74] The study arms are treated as independent. This as a limitation of this model, which may inflate the uncertainty in the estimates. Models and code that can incorporate non-independent data (measured on the same patients) with ordered categories are not readily available.

Studies that only report the number of individuals in collapsed categories, for example the number of individuals with no cancer, non-clinically significant cancer (Gleason 3+3) and clinically significant cancer (Gleason > 3+3) provide information only on the odds ratio of being classified in the first 2 categories (no cancer, non-clinically significant cancer). The model has been adapted to allow these studies to be included. However, they provide only limited information to the network compared to studies that report a finer breakdown of Gleason scores.

Models were fitted in WinBUGS 1.4.3.[75] Cognitive fusion prostate biopsy was chosen as the reference intervention, and 'no cancer' as the reference category. Full details of the model and WinBUGS code are given in Appendix 2.

The relative effects produced by the model are the odds ratios for being classified in category r, instead of category 1 ('no cancer'), using intervention X (software fusion, systematic biopsy or a combination of software/cognitive fusion with systematic biopsy), compared to cognitive fusion biopsy. Interpretation of these relative effects is complex since it relates to both a reference treatment and reference category. To aid interpretation, absolute probabilities of being classified in each category using each intervention are also reported. Details of how these are calculated are given in Appendix 2.

Analyses are presented assuming all software fusion devices share a common effect, that is they all have the same odds ratio compared to cognitive fusion biopsy (Model 1a) and assuming individual device effects (Model 1b).

4.2.3 Cancer detection NMA models

The odds ratios of cancer detection for different biopsy methods compared to each other were also pooled. The number of cancers detected were modelled using the NMA model for binomial data with a logit link described in NICE TSD 2,[71] fitted in R[72] using the package *gemtc*.[76]

Model convergence was assessed through inspection of Gelman-Rubin diagnostic plots. Both fixed-effect and random-effect models were fitted to the data. Non-informative prior distributions were used for all effect parameters and a Uniform(0,5) prior distribution was selected for the between-study standard deviation in random-effects models.[71] Model fit was assessed through mean total residual deviance and inspection of residual deviance contribution for each study arm. Heterogeneity was assessed by inspecting the size of the between-study standard deviation and its 95% credible interval, and by comparing the Deviance Information Criteria (DIC) for fixed-effect and random-effects models. Where DIC differed by less than three points the simplest model (fixed effect) was chosen. Consistency between direct and indirect evidence was assessed by fitting an unrelated mean effects model and where that suggested potential inconsistency, further investigation of the location of inconsistency was carried out by fitting node-split models.[77]

4.2.3.1 Any cancer detection NMA

The odds ratios of detecting any PCa (both clinically significant and non-clinically significant, i.e. Gleason ≥ 3+3) for different biopsy methods compared to each other were pooled. Analyses are presented assuming all software fusion devices share a common effect (Model 2a) and for individual device effects (Model 2b).

4.2.3.2 Clinically significant cancer detection NMA model

The odds ratios of detecting clinically significant PCa (Gleason > 3+3) for different biopsy methods compared to each other were also pooled, for studies that reported it. Analyses are presented assuming all software fusion devices share a common effect (Model 3a) and for individual device effects (Model 3b).

4.2.4 Narrative synthesis

Results of studies that were not eligible for inclusion in the NMAs, and results of all studies reporting protocol-specified outcomes other than diagnostic accuracy, were synthesised narratively following published guidelines.[78]

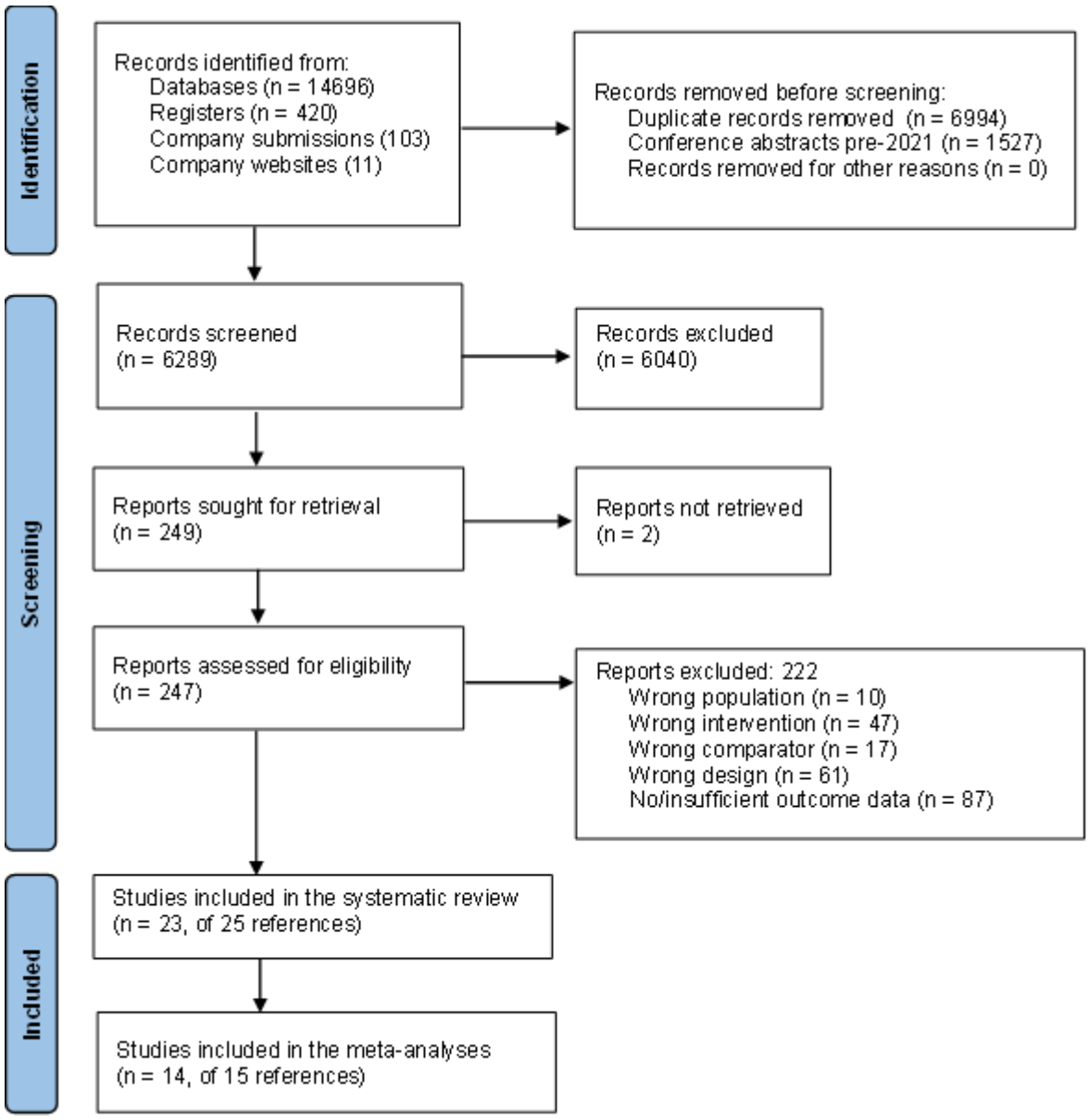
Outcomes were presented following the order listed in the protocol, then by comparison. Effect estimates, including metrics, measures of variance, statistical significance (at conventional threshold of p=0.05), and direction of effect were presented narratively and/or in tables at patient-level, unless only data per lesion could be extracted. Studies were grouped based on direction of effect and statistical significance. Where not reported, outcomes including detection rates, test positive rates and biopsy positivity rates were imputed. No formal statistical methods were used to assess heterogeneity. Results were narratively compared with the meta-analyses, and limitations of the evidence (e.g. inconsistency, risk of bias) informed findings summaries and conclusions.

4.3 Quantity and quality of evidence

Figure 2 presents an overview of the study selection process in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. The literature searches identified a total of 6289 unique records. After title and abstract screening, 247 references were retrieved and a total of 23 unique studies were included in the systematic review.[31, 79-101] Fourteen studies were included in the quantitative synthesis, [31, 79, 80, 82, 84, 86-88, 92-94, 96, 97, 99] while nine studies were included in the narrative synthesis only.[81, 83, 85, 89-91, 95, 98, 100, 101]

Evidence was included for all software fusion technologies specified in the scope and protocol (all versions) except for Fusion Bx (Focal Healthcare) and ExactVu (ExactImaging). A brief summary of the evidence for Fusion Bx and ExactVu that was considered for inclusion and ultimately excluded is presented in Appendix 3. A list of studies excluded from the systematic review, grouped by reason for exclusion, is reported in Appendix 3.

Figure 2. Study selection process (PRISMA flow diagram)



4.3.1 Description of studies included in the systematic review of diagnostic accuracy and clinical effectiveness

Table 4 presents the characteristics of the 23 studies included in the systematic review. The majority of studies were conducted in Europe, including Belgium (1 study),[94] Croatia (1 study),[86] France (3 studies),[92, 93, 98] Germany (3 studies),[83, 95, 101] Italy (3 studies),[81, 89, 101] the Netherlands (2 studies),[31, 80] and the UK (1 study).[95] Five studies were conducted in the USA.[87, 88, 90, 96, 97] The remaining studies were conducted in Australia (2 studies),[95, 100] China (1 study),[85] India (1 study),[79] Iran (1 study),[82] Mexico (1 study),[99] Saudi Arabia (1 study).[84] Two studies were conducted in more than one country.[91, 95, 101]

Twelve studies compared software fusion against cognitive fusion; of those, three used a within-patient comparison design (where participants underwent biopsy with both software fusion and cognitive fusion within the same session),[88, 93, 97] and nine compared separate cohorts who received either software fusion or cognitive fusion biopsy (between-patient design).[31, 82, 85, 89, 90, 95, 98, 100, 101] Three studies compared two or more software fusion software against one another.[79, 86, 99] Five studies compared software fusion against systematic biopsy,[80, 87, 92, 94, 96] and three studies compared cognitive fusion with systematic biopsy.[79, 86, 99]

Three RCTs were included; of those, two compared software fusion against cognitive fusion,[31, 82] and one compared three software fusion devices.[83] All other studies were non-randomised trials or observational; of those, four studies used a retrospective design.[85, 90, 100, 101]

The following software fusion technologies were evaluated: Artemis (5 studies),[82, 84, 88, 96 97] Biojet (4 studies),[81, 83, 84, 89] BiopSee (2 studies),[31, 95] BK (2 studies, referred to as Predictive Fusion Software in one study[85] and MIM fusion software in another),[100] iSR'obot Mona Lisa (1 study),[101] and Uronav (1 study).[90] One study evaluated KOELIS Trinity,[83] and six studies evaluated, KOELIS Urostation, an earlier version of the software which used a third party ultrasound.[80, 81, 92-94, 98]

Table 5 maps the evidence by software fusion technology, biopsy route, anaesthesia method and registration method, and highlights a number of limitations in reporting and gaps in the evidence. Of the 20 studies that evaluated a software fusion technology, seven studies used software fusion for a transperineal biopsy, [31, 84, 85, 89, 95, 100, 101] and there was no evidence for Artemis, Koelis and Uronav used in the context of a transperineal biopsy. Biopsee was only evaluated under general anaesthesia,[31, 95] and ten studies did not report their method of anaesthesia.[80, 81, 89, 91-94, 96, 98, 100] Image registration methods (rigid vs elastic) were not reported or could not be inferred in five studies.[84, 88, 96 89, 95]

Table 6 summarises the characteristics of the patients included in the included studies. Across all included studies, a total of 3733 patients who received software fusion and 2154 individuals who

underwent cognitive fusion were analysed and informed estimates of PCa detection. Where reported, the median age ranged from 62 to 73.1 years, median PSA levels ranged from 4.2ng/mL to 10.7ng/mL, and all patients had a PI-RADS or Likert score of 3 or more. Seven studies only included biopsy naïve patients,[79, 81, 82, 85, 88, 95, 98] four studies only included patients who received a repeat biopsy following one or more prior negative biopsies[31, 86, 94, 99] and eight studies included a mix of patients with no prior biopsy and individuals undergoing a repeat biopsy following a prior negative biopsy.[80, 83, 84, 87, 89-93, 101] Three studies included a subset of patients under active surveillance and reported separate results biopsy naïve and/or repeat biopsy with prior negative result.[96, 97, 100]

Table 4. Study characteristics of studies included in the systematic review

Study	Country	Design	N	Population investigated	MRI type	MRI magnet strength (T)	Software fusion technology	Comparison	Biopsy route	N of cores per lesion (targeted biopsy)[<]	Number of ROI targeted	N of cores (SB)	Anaesthesia	Definition of CsPCa	Definition of PCa
Software Fusion vs Cognitive Fusion: Prospective															
Cornud (2018)[93]	France	Prospective, within-patient	88	BN, RB	mpMRI	1.5	Urostation Touch (KOELIS)	SF vs CF[#]	TR	2	NR	NR	NR	NA	Gleason 3+3
Delongchamps (2013) [98]	France	Consecutive series, between patient	SF: 82 CF: 54	BN	mpMRI	1.5	Urostation Touch (KOELIS)[>]	SF vs CF vs SB	TR	≥2	NR	10-12	NR	Gleason ≥3+4	Gleason 3+3
FUTURE (2019) [31, 102]	Netherlands	RCT, between patient	SF: 79, CF: 78	RB	mpMRI	3	BiopSee (MedCom)	SF vs CF	SF: TP, CF: TR	Median (IQR). SF: 4 (3-5), CF: 3 (3-4)	All	NA	General/ spinal	Gleason ≥3+4	NR
Hansen (2018)[95]	UK, Germany, Australia	Prospective, between patient	SF: 395 CF: 176	BN	mpMRI	1.5 or 3	Biopsee (Medcom)	SF vs CF vs SB	TP	At least 2	All ROI	18-24[$]	General	Gleason ≥3+4	Gleason 3+3
Izadpanahi (2021)[82]	Iran	RCT, between patient	SF: 99 CF: 100	BN	mpMRI	3	Artemis (InnoMedicus Artemis)	SF vs CF, ±SB	TR	SF: 1-2 CF: 1-2	2	4	Local	Gleason ≥3+4, or 3+3 with ≥4mm core length	Gleason Score 3+3 with <4mm core length
PAIREDCAP (2019)[88]	USA	Prospective, within-patient	248	BN	mpMRI	3	Artemis (InnoMedicus Artemis)	SF vs CF vs SB	TR	SF: 3 CF: 3	1	12	Local	Gleason ≥3+4	Gleason ≥3+3
PROFUS (2014)[97]	USA	Prospective, within patient	101 (BN, RB)	BN, RB, AS	mpMRI	3	Artemis (InnoMedicus Artemis)	SF vs CF	TR	SF: 2 CF: 2	2	12[$]	Local	Gleason ≥3+4	Gleason 3+3
Stabile (2018)[89]	Italy	Prospective, between patient	SF: 157 CF: 87	BN, RB	mpMRI	1.5	Biojet	SF+SB vs CF+SB	SF: TP/TR CF: TR	Median (range) SF: 3 (2-3); CF: 2 (2-5)	All ROI	12	NR	Gleason ≥3+4	NR
Software Fusion vs Cognitive Fusion: Retrospective															
Kaufmann (2018)[91, 101]	Germany, Italy	Retrospective, between patient	SF: 191 CF: 87	BN, RB	mpMRI	3	iSR'obot Mona Lisa (Biobot Surgical)	SF vs CF	SF: TP CF: TR	4	1	14[$]	NR	Gleason ≥3+4	Gleason 3+3

Study	Country	Design	N	Population investigated	MRI type	MRI magnet strength (T)	Software fusion technology	Comparison	Biopsy route	N of cores per lesion (targeted biopsy)[<]	Number of ROI targeted	N of cores (SB)	Anaesthesia	Definition of CsPCa	Definition of PCa
Liang (2020)[85]	China	Retrospective, between patient	SF: 92 CF: 71	BN	bpMRI	3	Predictive Fusion Software (BK)	SF vs CF	TP	4	All ROI	NA	Local	Gleason ≥3+4	Gleason 3+3
Lockhart (2022)[100]	Australia	Retrospective, between patient	SF: 131 CF: 224	BN	mpMRI	3	MIM Fusion Software (with BK 3000 US)	SF+SB vs CF+SB	TP	NR[!]	NR	NR	NR	Gleason ≥3+4	NR
Monda (2018)[90]	USA	Retrospective; before and after study	SF: 348 CF: 162	BN, RB (+ve/-ve)	mpMRI	3	UroNav (Invivo Corporation)	SF vs CF vs SB	TR	NR	NR	12	NR	Gleason ≥3+4	Gleason 3+3
Software Fusion vs Software Fusion															
Ferriero (2022)[81]	Italy	Prospective cohort, between patient	Urostation: 103 Biojet: 232	BN	mpMRI	NR	Urostation (KOELIS) Biojet (Healthcare Supply Solutions Ltd)	SF vs SF	Urostation: TR; Biojet: NR	Median (IQR) **Unmatched** Urostation: 4 (4-6) Biojet: 6 (5-6) **Matched** Urostation: 4 (4-6) Biojet: 6 (4-6)	NR	NA	NR	Gleason ≥3+4	Gleason 3+3
Rabah (2021)[84]	Saudi Arabia	RCT, between patient	Artermis: 165 Biojet: 142	BN, RB[^]	mpMRI	NR	Artemis (InnoMedicus Artemis) Biojet (Healthcare Supply Solutions Ltd)	SF vs SF vs SB	Artemis: TR Biojet: TP	2-4	All ROI	12	Artemis: Local Biojet: General		

Study	Country	Design	N	Population investigated	MRI type	MRI magnet strength (T)	Software fusion technology	Comparison	Biopsy route	N of cores per lesion (targeted biopsy)<	Number of ROI targeted	N of cores (SB)	Anaesthesia	Definition of CsPCa	Definition of PCa
Sokolakis (2021)[83]	Germany	Prospective, between patient	Biojet: 20 Urnoav: 20 KOELIS Trinity: 20	BN, RB	mpMRI	3	Biojet (Healthcare Supply), UroNav (Phillips), KOELIS Trinity	SF vs SF	TR	2-3	All ROI	12$	Local	NR	Gleason 3+3
Software Fusion vs Systematic Biopsy vs Software Fusion and Systematic Biopsy															
Alberts (2018)[80]	The Netherlands	Prospective, within patient	48~	BN, RB	mpMRI	NR	Urostation (Koelis)	SF vs SB vs SF+SB	TR	2	All ROIs	12	NR	Gleason ≥3+4	Gleason 3+3
Albisinni (2018)[94]	Belgium	Prospective, within-patient	74	RB	mpMRI	3	Urostation (Koelis)	SF vs SB vs SF+SB	TR	2-4	1	12-14	NR	Gleason ≥3+ 4 and/or cancer core length ≥6mm (UCL)	NR
Filson (2016)[96]	USA	Prospective, within-patient	538 (PI-RADS ≥3, excl AS)	BN, RB, AS (not reported)	mpMRI	3	Artemis (InnoMedicus Artemis)	SF vs SB vs SF+SB	NR	1	NR	12	NR	Gleason ≥3+4	Gleason 3+3
Fourcade (2018)[92]	France	Prospective, within-patient	191	BN, RB	mpMRI	3	Urostation (Koelis)	SF vs SB vs SF+SB	TR	2-4	All ROI	12	NR	NR	NR
Wajswol (2020)[87]	USA	Prospective, within-patient	169 (PI-RADS ≥3)	BN, RB	mpMRI	3	UroNav (Phillips)	SF vs SB vs SF+SB	TP	4-6	All ROI	12	Local	Gleason ≥3+ 4	NR
Cognitive Fusion vs Systematic Biopsy vs Cognitive Fusion and Systematic Biopsy															
Gomez-Ortiz (2022)[99]	Mexico	Prospective, within-patient	111	RB	NR	1.5	N/A	CB vs SB vs CB+SB	TR	2-4	All ROI	12	NR	Gleason ≥3+4	Gleason 3+3
Kulis (2020)[86]	Croatia	Prospective, within-patient	63	RB	mpMRI	3	N/A	CB vs SB vs CB+SB	NR	6	Up to 2	12	Local£	Gleason ≥3+4	Gleason Score ≤6
Thangarasu (2021)[79]	India	Prospective, within—patient	75	BN	mpMRI	3	N/A	CB vs SB vs CB+SB	TR	2	All ROI	12	Local£	Gleason ≥3+ 4	NR

<all targeted biopsy methods, unless otherwise specified; # SB performed at operator's discretion (N patients NR); ^did not report whether a subset of AS patients were also included; ~ subset who received TB and SB; £'periprostatic block'; > Also compared with Esaote SF; $ SB performed but results were NR and did not inform detection comparisons between targered biopsies; ! SF+SB: mean 21 (range 12-33), CF+SB: 26 (9-54).

BN: biopsy naive; RB: repeat biosy; AS: active surveillance; SF: software fusion; CF: cognitive fusion; SB: systematic biopsy; NA: not applicable; ROI, region of interest; IQR, interquartile range' PCA, prostate cancer; csPCA, clinically significant prostate cancer; PI-RADS, Prostate Imaging Reporting & Data System.

Table 5 Summary of characteristics of studies of software fusion included in the systematic review

Device	Author	N	Biopsy route			Anaesthesia			Image registration		
			TR	TP	NR	Local	General	NR	Rigid	Elastic	NR
Artemis (InnoMedicus)	Filson, 2016	538			X			X			X
	Izadpanahi, 2021	99	X			X				X	
	PAIRED CAP, 2019	248	X			X					X
	PROFUS, 2014	101	X			X				X	
	Rabah, 2021	165	X			X					X
	TOTAL	1151	**4**	**0**	**1**	**4**	**0**	**1**	**0**	**2**	3
BiopSee (Medcom)	FUTURE, 2019	79		X			X		X		
	Hansen, 2018	395		X			X				X
	TOTAL	**474**	**0**	**2**	**0**	**0**	**2**	**0**	**1**	**0**	**1**
BK	Liang, 2020	92		X		X			X		
	Lockhart, 2022	131		X				X	X		
	TOTAL	223	**0**	**2**	**0**	**1**	**0**	**1**	**2**	**0**	**0**
Koelis (earlier versions)	Alberts, 2018	48	X					X		X	
	Albisinni, 2018	74	X					X		X	
	Cornud, 2018	88	X					X		X	
	Delongchamps, 2013	82	X					X		X	
	Ferriero, 2022	103			X			X		X	
	Fourcade, 2018	191	X					X		X	
	TOTAL	**586**	**5**	**0**	**1**	**0**	**0**	**6**	**0**	**6**	**0**
BioJet	Ferriero, 2022	232			X			X	X		
	Rabah, 2021	142		X			X				X
	Sokolakis 2021	20	X			X			X		
	Stabile, 2018	157	X	X				X			X
	TOTAL	**551**	**2**	**2**	**1**	**1**	**1**	**2**	**2**	**0**	**2**
iSR'obot Mona Lisa	Kaufmann, 2018	191		X				X		X	
	TOTAL	**191**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**1**	**0**
UroNav	Monda, 2018	348	X					X		X	
	Sokolakis, 2021	20	X			X			X		
	Wajswol, 2020	169		X		X			X	X	
	TOTAL	**537**	**2**	**1**	**0**	**2**	**0**	**1**	**2**	**2**	**0**
Koelis Trinity	Sokolakis, 2021	20	X			X				X	
	TOTAL	**20**	**1**	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**0**

TR, transrectal; TP, transperineal; NR, not reported

Table 6 Study and population characteristics of studies included in the systematic review

Study	N	Population investigated	Recruitment criteria	Age (yr)	PSA (ng/ml)	PI-RADS version	PI-RADS score			Lesion location
							3	4	5	
Software Fusion vs Cognitive Fusion										
Cornud (2018)[93]	88	BN, RB	PI-RADS ≥ 3*	Med (IQR) 63 (60-69)	Med (IQR) 8.2 (6.0-10.9)	NR	NR			NR
Delongchamps (2013)[103>]	mpMRI +ve SF: 82 CF: 54	BN	PSA ≥ 4ng/ml, and/or suspicious DRE	Mean (SD)+ SF: 64.5 (7.9) CF: 62.7 (7.4)	Mean (SD)+ SF: 8.3 (4.1) CF: 9 (3.9)	NR	NR			NR
FUTURE (2019)[31]	SF: 79 CF: 78	RB	Repeat SB (<4 yr), PSA≥4 (ng/ml) and/or suspicious DRE	Mean (SD) SF: 64.6 (6.9) CF: 66.5 (6.3)	Mean (SD) SF: 11.6 (9.0) CF: 11.0 (7.1)	v2	SF: 23 CF: 21	SF: 34 CF: 32	SF: 22 CF: 25	SF: 35 Post, 37 Ant CF: 46 Post, 25 Ant
Hansen (2018)[95]	PI-RADS ≥3 SF: 395 CF: 176	BN	PSA ≤30ng/mL, ≤79 years	Median (IQR)+ Centre 1 [SF]: 64 (57-69) Centre 2 [SF]: 65 (60-70) Centre 3 [CF]: 65 (60-70)	Median (IQR)+ Centre 1 [SF]: 6.6 (4.6-9.0) Centre 2 [SF]: 6.9 (5.2-9.1) Centre 3 [CF]: 5.9 (4.6-8.0)	v1-2	Centre 1 [SF]: 34, Centre 2 [SF]: 91, Centre 3 [CF]: 28	Centre 1 [SF]: 99, Centre 2 [SF]: 171, Centre 3 [CF]: 148		NR
Izadpanahi (2019)[82]	SF: 99; CF: 100	BN	PSA 2-10 ng/dL, PI-RADS≥3	Mean (SD) SF: 61.9 (7.4) CF: 61.9 (7.4)	Mean (SD) SF: 6.1 (1.3) CF: 5.9 (1.3)	v2	NR			NR
PAIREDCAP (2019)[31]	248	BN	Elevated PSA (serum PSA <25ng/mL) or abnormal DRE	Mean (SD) 65.5 (7.7)	Med (IQR) 6.2 (4.6-8.20	v2	56	91	101	Ant: 93
PROFUS (2014)[97]	125 (101 BN, RB)	BN, RB, AS	NR	NR (range) 65 (56.3-71.0)	NR (range)	v2	NR			Post: 140 Ant: 32

Study	N	Population investigated	Recruitment criteria	Age (yr)	PSA (ng/ml)	PI-RADS version	PI-RADS score			Lesion location
							3	4	5	
Stabile (2018)[89]	SF: 157 CF: 87	BN, RB	NR	Median (IQR) SF: 67 (61-73) CF: 62 (58-70)	Median (IQR) SF: 7.3 (5.2-10.5) CF: 6 (4-9)	NR	SF: 59, CF: 35	SF: 98 CF: 52		NR
Software Fusion vs Cognitive Fusion – Retrospective										
Kaufman (2018)[91, 101]	SF: 191 CF: 87	BN, RB	Rising and/or persistently elevated PSA	Median (IQR): 69.0 (63.0-74.0)	Median (IQR): 8.0 (5.87-12.0)	v2	NR			NR
Liang (2020)[85]	SF: 92 CF: 71	BN	PSA level of ≤20 ng/mL	Mean (SD) SF: 69.17 (9.18) CF: 67.59 (8.45)	Median (IQR) SF: 8.03 (0.66–19.78) CF: 7.66 (0.67–18.81)	v2	NR			NR
Lockhart (2022)[100]	Total: 355 (SF: 131, CF: 224); BN only: 283 (SF: 97; CF: 186)	BN, AS	NR	Mean (range) SF: 65 (41-80) CF: 66.6 (44-85)	Mean SF: 5.8 CF: 7.64	NR	NR			NR
Monda (2018)[90]	SF: 348 CF: 162	BN, RB (+ve/-ve)	NR	Mean (SD) SF: 65.0 (7.2) CF: 63.9 (7.8)	Mean (SD) 7.8 (7.8) 7.9 (7.8)	v2	NR			NR
Software Fusion vs Software Fusion										
Ferriero (2022)[81]	Unmatched Urostation: 103 Biojet: 232 Matched: Urostation: 83 Biojet: 83	BN	PI-RADS ≥3	Median (IQR) Unmatched Urostation: 67 (59, 72) Biojet: 60 (65, 75) Matched Urostation: 69 (60, 72) Biojet: 65 (61, 71)	Median (IQR) Unmatched Urostation: 7 (4.9, 10.3) Biojet: 6.5 (5, 5.95) Matched Urostation: 7 (4.9, 10.3) Biojet: 6.6 (5, 10)	NR	Unmatched: Urostation: 21 Biojet: 52. Matched: Urostation: 50 Biojet: 19.	Unmatched: Urostation: 55 Biojet: 108 Matched: Urostation: 26 Biojet: 43	Unmatched: Urostation:27 Biojet: 51 Matched: Urostation: 15 Biojet: 21	NR

Study	N	Population investigated	Recruitment criteria	Age (yr)	PSA (ng/ml)	PI-RADS version	PI-RADS score			Lesion location
							3	4	5	
Rabah (2021)[84]	Artermis: 165 Biojet: 142	BN, RB	PI-RADS ≥3, and PSA ≥3.5ng/ml or abnormal DRE	Mean (SD) Artemis: 65.1 (7.8) Biojet: 65 (8.5)	Mean (SD) Artemis: 14.2 (5) Biojet: 13.7 (25.9)	v2	Artemis: 35 Biojet: 30	Artemis: 19 Biojet: 25	Artemis: 16 Biojet: 20	NR
Sokolakis (2021)[83]	Biojet: 20 Urnoav: 20 KOELIS Trinity: 20	BN, RB	PI-RADS ≥3	Median (IQR) Biojet: 66 (61, 67) Uronav: 64 (61, 74) Trinity: 64 (62, 67)	Median (IQR) Biojet: 8 (6,9) Uronav: 6 (5,8) Trinirty: 7 (5,8)	v2	Biojet: 4 Uronav: 6 Trinity: 6	Biojet: 12 Uronav: 7 Trinity: 9	Biojet: 4 Uronav: 7 Trinity: 5	NR
Software Fusion vs Systematic Biopsy vs Software Fusion and Systematic Biopsy										
Alberts (2018)[80]	48 (who received TB and SB)	BN, RB	PI-RADS ≥3, and PSA ≥3.5ng/ml	Median (IQR)+ 73.1 (72.4-73.8)	Median (IQR)+ 4.2 (3.4-5.8)	NR	NR			NR
Albisinni (2018)[91]	74	RB	NR	Median (IQR) 65 (62-69)	Median (IQR) 9.27 (6.84-13.4)	v2	NR			NR
Filson (2016)[96]	538 (PI-RADS ≥3, excl AS)	BN, RB, AS (not reported)	Elevated PSA or abnormal DRE	Median (IQR) BN: 64.4 (58.5–69.4) RB: 65.7 (59.3–70.2)	Median (IQR) BN: 5.8 (4.4–8.1) RB: 7.6 (5.0–11.5)	v2	BN: 129 RB:148	BN: 109 RB: 87	BN: 35 RB: 30	Anterior: BN: 148 RB: 100
Fourcade (2018)[92]	191	BN, RB	PSA >4ng/mL and abnormal DRE	Median (range) 66 (47-80)	Mean (range) 9 (0.7-48)	v2	NR			
Wajswol (2020)[87]	169 (PI-RADS ≥3)	BN, RB	PI-RADS ≥2 (visible lesion), PSA >2.5ng/mL	Median (range) 67.5 (44-89)	Median (range) 8.25 (1.4-103.8)	v2	26	76	67	NR
Cognitive Fusion vs Systematic Biopsy vs Cognitive Fusion and Systematic Biopsy										
Gomez-Ortiz (2022)[99]	111	RB	PI-RADS ≥3	Mean (SD) 66.27 (6.85)	Median (IQR) 9.9 (1.21-26)	2	NR			NR

Study	N	Population investigated	Recruitment criteria	Age (yr)	PSA (ng/ml)	PI-RADS version	PI-RADS score			Lesion location
							3	4	5	
Kulis (2020)[86]	63	RB	PI-RADS ≥3, PSA >4ng/mL	Median (range) 67 (57-84)	Median (range) 10.70 (4.86-64.00)	v2	12	35	16	Central: 42 Peripheral: 9 Apical: 9 Anterior: 3
Thangarasu (2021)[79]	75	BN	PI-RADS ≥3, serum PSA > 4 and ≤20 ng/mL, suspected ≤T2 stage on rectal examination	Mean (SD) 66.31 (7.9)	Median (NR) 10.6 (4.5-20)	v2	42	23	10	NR

*PI-RADS version 1. > Also compared with Esaote Software Fusion, + not specific to the population of interest
BN: biopsy naive; RB: repeat biosy; AS: active surveillance; SD: standard deviation; NR: not reported; IQR: interquartile range; PSA, prostate specific antigen; PI-RADS: Prostate Imaging Reporting & Data System; SB: systematic biopsy; SF: software fusion; CF: cognitive fusion; SB: systematic biopsy

4.3.2 Quality of included studies

Results of the quality and applicability assessment are reported in Table 7, and further details on the rationale for decisions are reported in Appendix 5. All studies were at high risk of bias for at least one of the following domains: patient selection, index test, reference standard and flow and timing. Eight studies were at high risk of patient selection bias; all were non-randomised comparisons.[81, 83, 89, 90, 95, 98, 100, 101] Three studies were at unclear risk of selection bias; [84, 85 31]including the two randomised controlled trials,[31, 84] and all other studies were at low risk of selection bias. Nine studies had a high risk of bias related to the comparison of index tests, [31, 81{Wegelin, 2019 #1358, 84, 88, 89, 93, 97, 98, 101] and all other 14 studies were at low risk of bias for this domain.

Twenty studies were at high risk of bias associated with the reference standard.[31, 79-87, 89, 90, 92, 94-96, 98-101] For between-patient comparisons, this was primarily due to the fact that total cancer positive cases in each study arm or cohort were derived from different biopsy methods; in within-patient comparisons, as all biopsy methods were performed within the same examination, it was not feasible for studies to truly blind operators from tracks of preceding biopsy methods (true blinding would several biopsy sessions per patient, which would be unethical). Participants in all within-patient comparison studies received software fusion, cognitive fusion and/or systematic biopsy within the same examination; the order in which the different biopsy methods were implemented varied where reported, therefore the overall direction of bias due to the lack of operator blinding could not be determined.

Of the 15 studies that compared software fusion with cognitive fusion or with another software fusion device,[88, 93, 98 31, 81, 82, 95, 100 83-85, 89, 90, 97, 101] seven did not use systematic biopsy or include systematic biopsy results as part of a reference standard test.[31, 81 83-85, 97, 101] Of the studies that included systematic biopsy as part of a reference standard test, only one reported blinding the systematic biopsy operator to the MRI report.[88] This is an important design limitation, since knowledge of the MRI report may have influenced the placing of systematic biopsy cores. Clinical advisers to the EAG confirmed that lack of blinding to MRI reports may have improved the accuracy of systematic biopsies relative to targeted biopsies. Therefore, for most of the evidence for systematic biopsy included in this review, there is a risk that the detection of prostate cancer from systematic biopsy may have been overestimated compared with true random, standard systematic biopsy. This said, the lack of blinding to MRI report when using systematic biopsy concomitant with targeted biopsy is reflective of current practice. Blinding of the histopathologists who analysed the biopsy samples was generally not reported, and none of studies used template-guided mapping biopsy. Two studies were at high risk of bias due to missing outcomes data (flow and timing domain),[93, 95] and all other studies were at low risk of bias for this domain.

Three studies raised no concerns about their applicability to the review question.[88 79, 82] Five studies included a population that was deemed not applicable to the review question (high concern),[31, 86, 90, 94, 99] and five included a significant proportion (approximately half) of patients undergoing repeat biopsy following a prior negative biopsy.[87, 89, 92, 93, 101] Although patients with a prior negative biopsy were eligible in this systematic review, clinical advisers to the EAG noted that they made up only a minority (approximately under 10%) of the total population undergoing targeted biopsy who are not under active surveillance. All other studies included mostly biopsy naïve patients and had a population that was considered broadly representative. Five studies used an intervention that was not considered applicable to the review question,[31, 84, 89, 95, 101] primarily due to the use of general anaesthesia in all procedures. Clinical advisers to the EAG noted that general anaesthesia is normally only used in a minority of patients, although it may facilitate biopsy targeting due to the lack of patient movement. The applicability of software fusion was uncertain in ten studies.[80, 81, 90, 92-94, 96, 98-100] In four cases, this was due to insufficient reporting about biopsy routes and anaesthesia methods,[96 90, 99, 100] and in six studies, a KOELIS device with no integrated ultrasound was evaluated, and the applicability of their results to KOELIS Trinity was uncertain.[80, 81, 92-94, 98] Following request for further information from the EAG, the company did not clarify or provide evidence that the diagnostic accuracy of older versions of KOELIS was equivalent to KOELIS Trinity. Seven studies raised concerns about the applicability of the reference standard test.[31, 81, 83-85, 95, 101]

Table 7 Risk of bias for relative diagnostic accuracy estimates and applicability assessment of studies included in the systematic review

Study	Tests	Reference standard (or tests to estimate total confirmed cases)	Risk of bias (QUADAS-C)				Applicability concerns (QUADAS-2)		
			P	I	R	FT	P	I	R
Alberts 2018[80]	SF (Koelis Urostation)	SF+SB	✓	✓	✗	✓	✓	?	✓
	SB								
Albisinni 2018[94]	SF (Koelis Urostation)	SF+SB	✓	✓	✗	✓	✗	?	✓
	SB								
Cornud 2018[93]	SB (Urostation Touch, Koelis)	SF+CF±SB	✓	✗	✓	✗	?	✓	?
	CF							?	
Delongchamps 2013[98]	SF (Urostation Touch, Koelis)	SF+SB	✗	✗	✗	✓	✓	?	✓
	CF	CF+SB						✓	
Elkhoury 2019[88] (PAIREDCAP)	SF (Artemis)	SF+CF+SB	✓	✗	✓	✓	✓	✓	✓
	CF								
Ferriero 2022[81]	SF (Urostation, KOELIS)	SF (Urostation, KOELIS)	✗	✗	✗	✓	✓	?	✗
	SF (Biojet)	SF (Biojet)							
Filson 2016[96]	SF (Artemis)	SF+SB	✓	✓	✗	✓	✓	?	✓
	SB								

Fourcade 2018[92]	SF (Koelis Urostation) SB	SF+SB	✓	✓	✗	✓	?	?	✓
FUTURE[31]	SF (Biopsee) CF	SF CF	?	✗	✗	✓	✗	✗ ✓	✗
Gomez-Ortiz 2022[99]	CF SB	CF+SB	✓	✓	✗	✓	✗	?	✓
Hansen 2018[95]	SF (Biopsee) CF	SF+SB CF+SB	✗	✓	✗	✗	✓	✗	✗
Izadpanahi 2021[82]	SF (Artemis)+SB CF+SB	SF+SB CF+SB	✓	✓	✗	✓	✓	✓	✓
Kaufmann 2018[91, 101]	SF (iSR'obot Mona Lisa) CF	SF CF	✗	✗	✗	✓	?	✗ ✓	✗
Kulis 2020[86]	CF SB	CF+SB	✓	✓	✗	✓	✗	✓	✓
Liang 2020[85]	SF (BK) CF	SF CF	?	✓	✗	✓	✓	✓	✗
Lockhart 2022[100]	SF (BK/MIM) CF	SF+SB CF+SB	✗	✓	✗	✓	✓	?	✓
Monda 2018[90]	SF (Uronav) CF	SF+SB CF+SB	✗	✓	✗	✓	✗	?	✓
PROFUS[97]	SF (Artemis) CF	SF+CF	✓	✗	?	✓	✓	✓	✗
Rabah 2021[84]	SF (Biojet) SF (Artemis)	SF (Biojet) SF (Artemis)	?	✗	✗	✓	✓	✗ ✓	✗
Sokolakis 2021[83]	SF (Biojet) SF (Uronav) SF (Koelis Trinity)	SF (Biojet) SF (Uronav) SF (Koelis Trinity)	✗	✓	✗	✓	✓	✓	✗
Stabile 2018[89]	SF (Biojet) CF	SF+SB CF+SB	✗	✗	✗	✓	?	✗	✓
Thangarasu 2021[79]	CF SB	CF+SB	✓	✓	✗	✓	✓	✓	✓
Wasjwol 2020[87]	SF (Uronav) SB	SF+SB	✓	✓	✗	✓	?	✓	✓

P: patient selection; I : index test; R : reference standard/test(s) used to derive overall test positive rates; FT : flow and timing; SF : software fusion; CF : cognitive fusion; TP: transperineal; TR: transrectal; GA: general anaesthesia; LA: local anaesthesia; NR: not reported

✓ indicates low risk; ✗ indicates high risk; ? indicates unclear risk

4.4 Diagnostic accuracy results

This section presents the evidence included in the meta-analyses and structure of the networks of evidence (Section 4.4.1), the results of the network meta-analyses (Section 4.4.2), and results of studies not included in the meta-analyses (Section 4.4.3).

4.4.1 Studies included in the meta-analysis and network structure

4.4.1.1 Model 1a: Multinomial synthesis model (base-case)

Thirteen studies identified by the systematic review with data suitable for inclusion in the NMA are presented in Table 8 and form the network in Figure 3. Rabah (2021)[84] is excluded as it compared two software fusion devices, assumed to have identical effects, and therefore does not contribute to the analysis. The multinomial synthesis model was used to synthesise comparative information on the probabilities of being classified at the various ISUP grades of PCa (Section 4.2.2). Resulting estimates are then used in the base-case economic model.

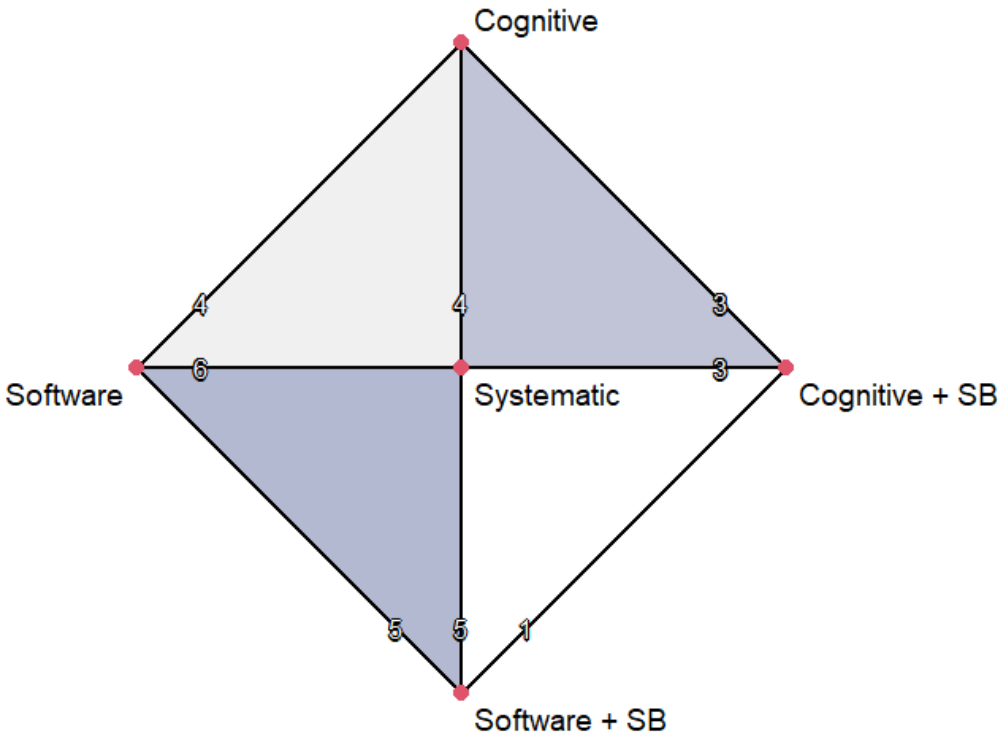
Due to data sparseness, we assumed that there is no difference in relative effects of the various software fusion biopsy devices compared to cognitive biopsy and only fixed-effect models could be fitted. This assumption is supported by the limited direct evidence comparing different fusion devices and clinical advice to the EAG. However, the different costs of each device will still be taken into account in the economic model. This assumption will be relaxed in an additional analysis (Model 1b, Section 4.4.1.2).

Table 8 Data for multinomial synthesis model.

	Biopsy type			number of patients			Category 1 No cancer			Category 2 ISUP grade 1			Category 3 ISUP grade 2			Category 4 ISUP grade 3			Category 5 ISUP grades 4-5		
Study	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**
PAIREDCAP (2019)[88]	CF	SB	Artemis	248	248	248	94	52	71	38	46	43	52	87	70	39	37	40	25	26	24
Izadpanahi (2021)[82]	CF + SB	Artemis + SB	NA	100	99	NA	69	55	NA	19	25	NA	6	13	NA	5	3	NA	1	3	NA
Wajswol (2020)[87]	SB	Uronav	Uronav + SB	169	169	169	53	49	36	116	120	133	NA	NA	NA	NA	NA	NA	NA	NA	NA
Thangarasu (2021)[79]	CF	SB	CF + SB	75	75	75	41	35	32	34	40	43	NA	NA	NA	NA	NA	NA	NA	NA	NA
Kulis (2020)[86]	CF	SB	CF + SB	63	63	63	30	33	25	33	30	38	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cornud (2018)[93]	CF	Urostation	NA	88	88	NA	57	48	NA	31	40	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
FUTURE (2019)[31]	CF	BiopSee	NA	78	79	NA	44	40	NA	8	12	NA	26	27	NA	NA	NA	NA	NA	NA	NA
PROFUS (2014)[97]	CF	Artemis	NA	125	125	NA	85	80	NA	16	16	NA	24	29	NA	NA	NA	NA	NA	NA	NA
Albisinni (2018)[94]	SB	Urostation	Urostation + SB	74	74	74	41	39	32	12	10	13	21	25	29	NA	NA	NA	NA	NA	NA
Fourcade (2018)[92]	SB	Urostation	Urostation + SB	191	191	191	103	106	85	36	25	34	52	60	72	NA	NA	NA	NA	NA	NA
Gomez-Ortiz (2022)[99]	CF	SB	CF + SB	111	111	111	69	81	65	19	9	20	23	21	26	NA	NA	NA	NA	NA	NA
*Rabah (2021)[84]	Artemis	Biojet	NA	165	142	NA	117	78	NA	27	18	NA	21	46	NA	NA	NA	NA	NA	NA	NA
Alberts (2018)[80]	SB	Urostation	Urostation + SB	48	48	48	23	20	16	11	11	13	10	13	13	4	4	6	NA	NA	NA
Filson (2016)[96]	SB	Artemis	Artemis + SB	538	538	538	294	310	252	114	68	100	74	81	92	56	79	94	NA	NA	NA

Studies are ordered by reported ISUP grade breakdown. Studies not reporting all ISUP breakdown, report data on the total number of patients classified at that ISUP grade or higher.
* Study only included in analyses with individual device effects as it compares two software fusion devices;
SF: software fusion; CF: cognitive fusion; SB: systematic biopsy; NA: not available/not applicable; ISUP: International Society of Urological Pathology

Figure 3 Network of biopsy types compared, under the assumption of a common effect for different software fusion devices.



Lines represent comparisons made in studies, numbers on the lines show how many studies included that comparison and shaded areas represent multi-arm studies.
Abbreviations: SB, systematic biopsy.

Although the network in Figure 3 is fully connected (there is a path connecting every intervention to every other), not all studies reported the breakdown of cancers detected by ISUP grade (see Table 8). This resulted in a *de facto* disconnect in the network for comparisons of cognitive fusion + SB and software fusion + SB for ISUP grades greater than 2. Relative effects comparing disconnected components of the network cannot be estimated and are reported separately.

Calculating absolute probabilities

As noted in Section 4.2.2, odds ratios estimated from this model are hard to interpret. We will therefore also present results on the absolute probability scale to aid interpretation. To calculate the absolute probabilities of being classified in each category using each intervention, we need to assume a set of underlying baseline probabilities of being classified in each category on one of the included interventions. For ease of interpretation, in this section these underlying baseline probabilities will be assumed to be fixed, i.e. to have no uncertainty. All other probabilities are then obtained by applying the estimated odds ratios to these probabilities, as described in Appendix 2. These baseline probabilities should be as representative as possible of the population of interest. A targeted review was carried out to determine a good source of evidence on these probabilities (Section 4.7.1.2).

The two studies with the largest sample size that were identified and deemed most representative of NHS practice were considered as a source of evidence for the baseline probabilities: Filson (2016)[96] and PAIREDCAP (2019).[88]

Two subgroups of patients are of interest: biopsy naïve patients and those undergoing a repeat biopsy after a negative result. Filson (2016)[96] reported probabilities for these two subgroups separately and for two interventions of interest, software fusion using Artemis and combined software fusion (Artemis) with systematic biopsy, allowing the same source of baseline probabilities to be used for both disconnected components of the network.

However, Filson (2016)[96] does not report separate data for ISUP grades 3 and 4-5, as required by the model. We approximated the probabilities of patients being in grade 3 and 4-5 by splitting the combined patients according to the proportions in each category reported in PAIREDCAP (2019)[88] (approximately 60/40).

In a sensitivity analysis for the subgroup of biopsy-naïve patients, the distribution of test results from PAIREDCAP (2019)[88] (which only include biopsy naïve patients) was used to inform the baseline probabilities in the first part of the network. In the absence of other suitable sources of evidence, data on biopsy-naïve patients from Filson (2016)[96] will continue to inform the baseline probabilities in the combined biopsy (software/cognitive fusion plus systematic biopsy) network.

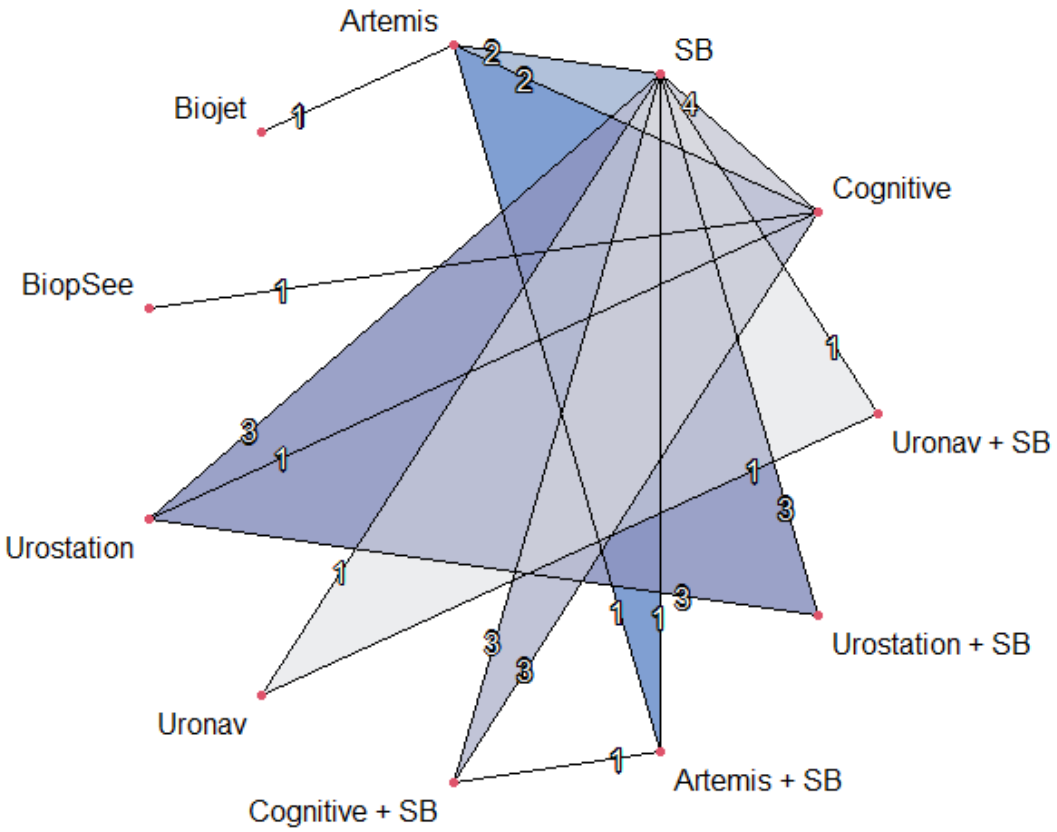
Absolute probabilities were therefore reported for:

- Subgroup of biopsy-naïve patients (based on Filson (2016)[96] biopsy-naïve data)
- Subgroup of previous negative biopsy patients (based on Filson (2016)[96] previous negative biopsy data)
- A sensitivity analysis using alternative baseline probabilities for the biopsy-naïve subgroup (based on biopsy-naïve data from PAIREDCAP (2019)[88] and Filson (2016)[96])

Results will be reported separately for comparisons of cognitive fusion, software fusion and systematic biopsy, and comparisons of combined cognitive/software fusion with systematic biopsy.

4.4.1.2 Model 1b: Multinomial synthesis model, individual device effects

Fourteen studies identified by the systematic review with data suitable for inclusion in the NMA are presented in Table 8 and form the network in Figure 4. The multinomial synthesis model was used to synthesise comparative information on the probabilities of being classified at the various ISUP grades of prostate cancer (Section 4.2.2).



Lines represent comparisons made in studies, numbers on the lines show how many studies included that comparison and shaded areas represent multi-arm studies.
Abbreviations: SB, systematic biopsy.

Figure 4 Network of biopsy types and devices compared.

Although the network in Figure 4 is fully connected (there is a path connecting every intervention to every other), not all studies reported the breakdown of cancers detected by ISUP grades. This resulted in a *de facto* disconnect in the network for comparisons of some devices for higher ISUP grades (see Table 8). Relative effects comparing disconnected components of the network cannot be estimated and are reported separately, where possible.

Calculating absolute probabilities

Absolute probabilities will be reported for

- Subgroup of biopsy-naïve patients (based on Filson (2016)[96] biopsy-naïve data)
- Subgroup of previous negative biopsy patients (based on Filson (2016)[96] previous negative biopsy data)

As many network components are disconnected for high ISUP grades in this analysis, absolute probabilities are only reported where they can be reliably obtained, which limits the interpretation of results.

4.4.1.3 Model 2a: Cancer detection

Data from the studies identified by the systematic review (Figure 3) were pooled in an NMA to compare the proportion of prostate cancers (clinically significant and non-clinically significant, i.e., Gleason ≥3+3) detected by the different biopsy strategies. Data were obtained by adding the relevant ISUP grades in Table 8, and are presented in Table 63 (Appendix 6).

In Model 2a we assumed that there is no difference in relative effects of the various software fusion biopsy devices compared to cognitive biopsy. This assumption in relaxed in Model 2b where the individual device effects are estimated. Both fixed- and random-effects models were considered.

4.4.1.4 Model 2b: Cancer detection, individual device effects

Data from the studies identified by the systematic review (Figure 4 and Table 63, Appendix 6) were pooled in an NMA to compare the proportion of prostate cancers (clinically significant and non-clinically significant) detected by the different biopsy strategies. Both fixed- and random-effects models were considered.

4.4.1.5 Model 3a: Clinically significant cancer detection

Data from the studies identified by the systematic review were pooled in an NMA to compare the proportion of clinically significant prostate cancers (Gleason > 3+3) detected by the different biopsy strategies. Only 10 studies reported the number of clinically significant cancers detected, obtained by adding the relevant ISUP grades in Table 8, and are presented in Table 64, Appendix 6. In addition, Rabah (2021)[84] is excluded as it compared two software fusion devices, assumed to have identical effects in Model 3a, and therefore does not contribute to this analysis. Nine studies were included in the network (Figure 12, Appendix 6). Both fixed- and random-effects models were considered.

4.4.1.6 Model 3b: Clinically significant cancer detection, individual device effects

Data from 10 studies reporting the number of clinically significant prostate cancers detected by the different biopsy strategies (Table 64 and Figure 13, Appendix 6) were pooled in an NMA. Both fixed- and random-effects models were considered.

4.4.2 Meta-analysis Results

4.4.2.1 Model 1a: Multinomial synthesis model (base-case)

Models were sampled for 100,000 iterations from 2 independent chains (50,000 iterations on each chain) after checking that convergence was achieved after a burn-in of 50,000 iterations.

Results from fitting Model 1a to the data in Table 8 (network in Figure 3) are presented in Table 9 One study (Gomez-Ortiz (2022)[99]) had a higher than expected contribution to the mean residual deviance (15 compared to its expected contribution of 6) but overall the model fitted the data well with a posterior mean of the residual deviance of 77.4, which is close to the 75 data points included.

Table 9 Odds ratios (median and 95% CrI) of being classified as ISUP grades1 to 4-5 compared to being categorised as having no cancer, for systematic biopsy and software fusion biopsy, compared to categorisations using cognitive fusion biopsy; and for software fusion plus systematic biopsy, compared to cognitive fusion plus systematic biopsy.

	Compared to cognitive fusion biopsy		**Compared to cognitive fusion + systematic biopsy**
ISUP grade	**SB**	**SF**	**SF + SB**
No cancer	REFERENCE		
1	1.57 (1.09, 2.26)	1.98 (1.28, 3.06)	1.20 (0.72, 1.99)
2	2.24 (1.45, 3.47)	1.34 (0.80, 2.25)	2.57 (0.95, 7.97)
3	1.40 (0.82, 2.38)	1.25 (0.66, 2.33)	0.66 (0.12, 2.92)
4-5	1.54 (0.83, 2.84)	1.58 (0.90, 2.77)	4.33 (0.45, 158.38)

SB: systematic biopsy; SF: software fusion biopsy; ISUP grade: International Society of Urological Pathology

Compared to cognitive fusion biopsy, there is evidence of higher odds of being categorised in ISUP grade 1 instead of no cancer when using software fusion (OR 1.98 95%CrI 1.28 to 3.06, Table 9). There is no evidence of more patients being categorised as ISUP 2, 3 or 4-5 instead of no cancer for software fusion biopsy compared to cognitive fusion biopsy (Table 9). More patients are categorised as having non-clinically significant cancer (ISUP grade 1) (OR 1.57 95%CrI 1.09 to 2.26) and as having a clinically significant cancer with ISUP grade 2 (OR 2.24 95%CrI 1.45 to 3.47), instead of having no cancer when using systematic biopsy compared to cognitive fusion biopsy. There is no clear evidence of more patients being categorised as ISUP 3 or 4-5 instead of no cancer for systematic biopsy compared to cognitive fusion biopsy (Table 9). However, we note the large uncertainty in all results, particularly for higher ISUP grades, due to limited data broken down by higher ISUP grades. As discussed in 4.3.2, most of the evidence for systematic biopsy was not blinded to MRI reports. This may have inflated the accuracy of systematic biopsy compared with software fusion and cognitive fusion.

Compared to cognitive fusion plus systematic biopsy, there is no clear evidence of more patients being categorised as having cancer (ISUP grades 1 to 4-5) instead of no cancer for software fusion plus systematic biopsy. However, we note the large uncertainty in all results, particularly for the highest category. This is due to few studies reporting data broken down by higher ISUP grades and the small number of patients categorised as ISUP 4-5 using any of the two biopsy types (Table 8).

Absolute probabilities of being classified as having no cancer or at different ISUP grades for the two subgroups of interest: biopsy-naïve patients and patients with a previous negative biopsy based on data from Filson (2016)[96] are presented for ease of interpretation. A sensitivity analysis for the biopsy-naïve subgroup is presented in Table 65, Appendix 6.

Absolute probabilities: Biopsy-naïve patients

Using baseline probabilities for software fusion biopsy and software fusion plus systematic biopsy from the biopsy-naïve subgroup in Filson (2016),[96] and applying the odds ratios in Table 9 the probabilities of being classified as having no cancer or at different ISUP grades are given in Table 10.

Table 10 Probabilities (median and 95%CrI) of being classified at different ISUP grades for biopsy-naïve patients.

	Artemis probabilities from Filson (2016)[96] biopsy-naïve data					**Artemis + SB probabilities from Filson (2016)[96] biopsy-naïve data**		
ISUP	**Cognitive**		**Systematic**		**Software***	**Cognitive + SB**		**Software + SB***
No cancer	0.55	(0.48, 0.62)	0.42	(0.37, 0.47)	0.47	0.41	(0.21, 0.56)	0.36
1	0.17	(0.13, 0.22)	0.21	(0.17, 0.25)	0.16	0.21	(0.10, 0.33)	0.22
2	0.12	(0.08, 0.16)	0.20	(0.16, 0.24)	0.20	0.10	(0.03, 0.23)	0.22
3	0.09	(0.06, 0.14)	0.10	(0.06, 0.15)	0.11	0.21	(0.06, 0.59)	0.12
4-5	0.06	(0.03, 0.10)	0.07	(0.04, 0.12)	0.06	0.02	(0.00, 0.18)	0.08

* Assumed underlying baseline probabilities

For biopsy-naïve patients, given a 47% probability of being classified as not having cancer with software fusion biopsy,[96] the probability of being classified as not having cancer is higher for patients undergoing cognitive biopsy (55% 95%CrI 48% to 62%). The probability of being classified as having non-clinically significant cancer (ISUP grade 1) is similar for patients undergoing software (16%) or cognitive fusion biopsy (17% 95%CrI 13% to 22%). Therefore, there is a lower probability of patients being classified at higher ISUP grades, particularly ISUP 2, with cognitive fusion biopsy compared to software fusion biopsy (Table 10). Probabilities for systematic biopsy are similar to those for software fusion biopsy for no cancer and all ISUP grades.

For a 36% probability of being classified as not having cancer with software fusion biopsy combined with systematic biopsy,[96] the probability of being classified as not having cancer is higher for patients undergoing cognitive biopsy combined with systematic biopsy (41% 95%CrI 21% to 56%). The probability of being classified as having non-clinically significant cancer (ISUP grade 1) is similar for patients undergoing software plus systematic (22%) or cognitive plus systematic biopsy (21% 95%CrI 10% to 33%). There is a higher probability of patients being classified at ISUP grade 2 with software plus systematic biopsy compared to cognitive plus systematic biopsy, and vice versa for ISUP grade 3 (Table 10). The proportion of patients classified at ISUP grades 4-5 are similar.

Absolute probabilities: Previous negative biopsy patients

Using baseline probabilities for software fusion biopsy and software fusion plus systematic biopsy from the subgroup of patients with a previous negative biopsy in Filson (2016),[96] and applying the odds ratios in Table 9 the probabilities of being classified as having no cancer or at different ISUP grades are given in Table 11.

Table 11 Probabilities (median and 95%CrI) of being classified at different ISUP grades for patients with a previous negative biopsy.

	Artemis probabilities from Filson (2016)[96] previous negative biopsy data				**Artemis + SB probabilities from Filson (2016)[96] previous negative biopsy data**	
ISUP	**Cognitive**	**Systematic**	**Software***	**Cognitive + SB**	**Software + SB***
No cancer	0.75 (0.69, 0.80)	0.64 (0.59, 0.69)	0.69	0.63 (0.38, 0.76)	0.58
1	0.08 (0.06, 0.11)	0.11 (0.09, 0.14)	0.09	0.13 (0.07, 0.21)	0.15
2	0.06 (0.04, 0.08)	0.11 (0.09, 0.13)	0.10	0.05 (0.02, 0.12)	0.12
3	0.06 (0.04, 0.10)	0.08 (0.05, 0.12)	0.08	0.14 (0.04, 0.47)	0.09
4-5	0.04 (0.02, 0.07)	0.05 (0.03, 0.09)	0.05	0.01 (0.00, 0.13)	0.06

* Assumed underlying baseline probabilities

For patients with a previous negative biopsy, given a 69% probability of being classified as not having cancer with software fusion biopsy,[96] the probability of being classified as not having cancer is higher for patients undergoing cognitive biopsy (75% 95%CrI 69% to 80%) but lower for patients undergoing systematic biopsy (64% 95%CrI 59% to 69%). As there is high probability that patients with a prior negative biopsy will again be classified as having no cancer with software fusion, cognitive fusion or systematic biopsy, the probabilities of being classified at different ISUP grades are small and similar across these biopsy strategies (Table 11).

The probabilities of being classified as not having cancer or ISUP grade 1 with software fusion biopsy combined with systematic biopsy and cognitive fusion biopsy combined with systematic biopsy are similar. However, there is a higher probability of being classified at ISUP grade 2 with software plus systematic biopsy compared to cognitive plus systematic biopsy, and vice versa for ISUP grade 3 (Table 10). The proportion of patients classified at ISUP grades 4-5 are similar.

4.4.2.2 Model 1b: Multinomial synthesis model, individual device effects

Models were sampled for 100,000 iterations from 2 independent chains (50,000 iterations on each chain) after checking that convergence was achieved after a burn-in of 50,000 iterations.

Results from fitting Model 1b to the data in Table 8 (network in Figure 4) are presented in Table 66, Appendix 6. One study (Gomez-Ortiz (2022)[99]) had a higher than expected contribution to the mean

residual deviance (16 compared to its expected contribution of 6). Other studies had deviances in the range expected, although the posterior mean of the residual deviance was 89.2, which is higher than the number points included (79). Often model fit can be poor when data are sparse as many parameters cannot be reliably estimated. However, more complex models, such as random-effects models, cannot be considered due this data sparseness. We would advise caution when interpreting the results from this model.

No odds ratios can be estimated for software fusion biopsy using Uronav or Uronav plus systematic biopsy since the only study comparing this device does not report details of classifications broken down by category (Table 8).

Compared to cognitive fusion biopsy, there is only evidence of higher odds of being categorised in ISUP grade 1 instead of no cancer when using systematic biopsy (OR 1.54 95%CrI 1.06 to 2.24, Table 66). There is some evidence that more patients are categorised as ISUP grade 2 instead of having no cancer when using systematic biopsy, Artemis or Urostation, compared to cognitive fusion biopsy. There is no clear evidence of more patients being categorised as ISUP 3 or 4-5 instead of no cancer for systematic biopsy or Artemis compared to cognitive fusion biopsy. No relative effects are estimable for the other devices and there is large uncertainty in all results.

Compared to cognitive fusion plus systematic biopsy, there is no clear evidence of more patients being categorised as having cancer (ISUP grades 1 to 4-5) instead of no cancer for Artemis or Urostation plus systematic biopsy. However, we note the large uncertainty in all results which led to some relative effects not being estimable (Table 66, Appendix 6).

Absolute probabilities of being classified as having no cancer or at different ISUP grades for the two subgroups of interest can only be reported where the odds ratios are estimable for all ISUP grades: Therefore, these will only be presented for cognitive fusion, systematic biopsy and software fusion using Artemis (assumed underlying baseline probabilities), and cognitive biopsy plus systematic biopsy and Artemis (assumed underlying baseline probabilities), for biopsy-naïve patients and patients with a previous negative biopsy based on data from Filson (2016)[96] (Table 67 and Table 68, Appendix 6).

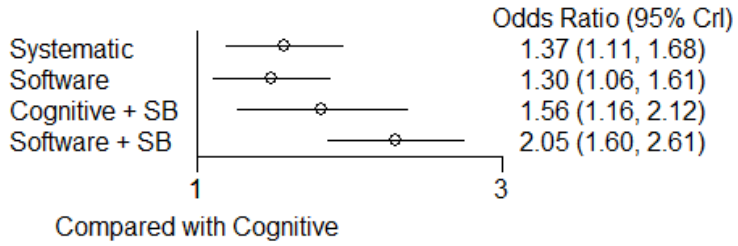
4.4.2.3 Model 2a: Cancer detection

Fixed- and random-effects models were fitted. Based on the model fit statistics (Table 69, Appendix 6) both the fixed- and random-effects models fitted the data well and differences in DIC were small. Therefore, the fixed-effect model was selected. The fixed-effect unrelated mean effects model suggested no evidence of inconsistency between direct and indirect evidence based on both the model fit statistics and deviance plots (Table 69 and Figure 14, Appendix 6).

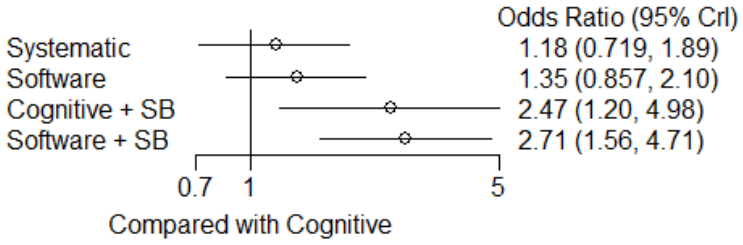
Results from fitting Model 2a to the data in Table 63, Appendix 6 (network in Figure 3) are presented in Figure 5 and all pairwise comparisons are reported in Table 70, Appendix 6.

Figure 5 Odds ratio of detection (median and 95%CrI) of cancer

a) Any cancer, fixed-effect NMA (model 2a)



b) Clinically significant cancer, random-effects NMA (model 3a)



CrI, credible interval, OR, odds ratio; SB, systematic biopsy

Compared to cognitive fusion biopsy, there is evidence that systematic biopsy, software fusion biopsy and the combination of cognitive or software fusion with systematic biopsy may detect more cancers than cognitive biopsy alone. Results for the random-effects model are presented as a sensitivity analysis in Table 70 and Figure 15, Appendix 6. As discussed above, 4.3.2, the accuracy of systematic biopsy may have been inflated due to study design limimations.

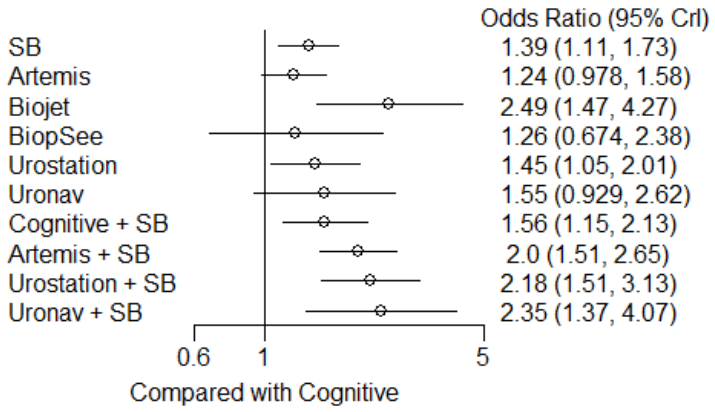
4.4.2.4 Model 2b: Cancer detection, individual device effects

Fixed- and random-effects models were fitted. Based on the model fit statistics (Table 69, Appendix 6) both the fixed- and random-effects models fitted the data well and differences in DIC were small. Therefore, the fixed-effect model was selected. The fixed-effect unrelated mean effects model suggested no evidence of inconsistency between direct and indirect evidence based on both the model fit statistics and deviance plots (Table 69 and Figure 14, Appendix 6).

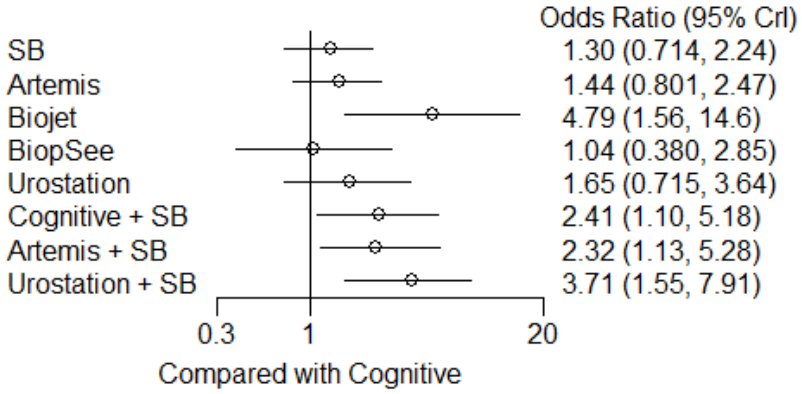
Results from fitting Model 2b to the data in Table 63, Appendix 6 (network in Figure 3) are presented in Figure 6 and all pairwise comparisons are reported in Table 71, Appendix 6.

Figure 6 Odds ratio of detection (median and 95%CrI) of cancer, individual device effects

a) Any cancer, fixed-effect NMA (model 2b)



b) Clinically significant cancer, random-effects NMA (model 3b)



CrI, credible interval, OR, odds ratio; SB, systematic biopsy

Compared to cognitive fusion biopsy, there is evidence that software fusion biopsy with Biojet, Urostation and Artemis, and Urostation, Uronav or cognitive biopsy combined with systematic biopsy, may detect more cancers.

Results for the random-effects model are presented as a sensitivity analysis in Table 71 and Figure 15, Appendix 6.

4.4.2.5 Model 3a: Clinically significant cancer detection

Fixed- and random-effects models were fitted. Based on the model fit statistics (Table 69, Appendix 6) the random-effects model had a better fit to the data and the difference in DIC was greater than 3. Therefore, the random-effects model was selected. The random-effects unrelated mean effects model suggested no evidence of inconsistency between direct and indirect evidence based on both the model fit statistics and deviance plots (Table 69 and Figure 14, Appendix 6).

Results from fitting Model 3a to the data in Table 64, Appendix 6 (network in Figure 12) are presented in Figure 5 and all pairwise comparisons are reported in Table 70, Appendix 6. The posterior median of the between-study heterogeneity standard deviation was 0.313 (95%CrI 0.132 to 0.634), which is moderate on the log odds ratio scale. The full posterior distribution of the between-study standard deviation is presented in Figure 15, Appendix 6.

Compared to cognitive fusion biopsy, there is no evidence that software fusion or systematic biopsy detect more clinically significant cancers. This may appear to contradict the results of Model 1a where it was found that more patients are categorised as having a cancer with ISUP grade 2 instead of having no cancer when using systematic biopsy, compared to cognitive biopsy. This result in Model 1a is largely driven by the data from PAIREDCAP which has fewer patients in ISUP 2 (52 patients) compared to systematic biopsy (87 patients) (Table 8). However, when considering ISUP 2 to 5 combined, additional information (via direct and indirect comparisons) is added from other studies reporting on csPCa as a combined category (and not specifically ISUP 2 alone), where differences are not apparent.

There is also evidence that adding systematic biopsy to cognitive or software fusion increases clinically significant cancer detection.

4.4.2.6 Model 3b: Clinically significant cancer detection, individual device effects

Fixed- and random-effects models were fitted. Based on the model fit statistics (Table 69, Appendix 6) the random-effects model had a better fit to the data and the difference in DIC was greater than 3. Therefore, the random-effects model was selected. The random-effects unrelated mean effects model suggested no evidence of inconsistency between direct and indirect evidence based on both the model fit statistics and deviance plots (Table 69 and Figure 14, Appendix 6).

Results from fitting Model 3b to the data in Table 64, Appendix 6 (network in Figure 12) are presented in Figure 6 and all pairwise comparisons are reported in Table 72, Appendix 6. The posterior median of the between-study heterogeneity standard deviation was 0.304 (95%CrI 0.048 to 0.769), which is similar to the posterior heterogeneity from Model 3a. This suggests there is moderate heterogeneity (log odds ratio scale) and that splitting the device effects did not explain the between-study variability. The full posterior distribution of the between-study standard deviation is presented in Figure 15, Appendix 6.

Compared to cognitive fusion biopsy, there is no evidence that software fusion with Artemis, BiopSee, Urostation, or systematic biopsy detect more clinically significant cancers. However, there is evidence that software fusion with Biojet or adding systematic biopsy to cognitive or software fusion with Artemis or Urostation increases clinically significant cancer detection.

4.4.3 Narrative synthesis of studies not included in the meta-analysis

Nine studies reported data on PCa detection but were not included in a meta-analysis, due to reasons specified in section 4.2.1.[81, 83, 85, 89-91, 95, 98, 100] None of these studies had a within-patient comparison, and none used a randomised comparison between software fusion and cognitive fusion or between two or more eligible software fusion technologies. Therefore, these studies were considered at higher risk of confounding compared with studies included in the NMA. This section presents a narrative summary of their results.

All nine studies reported a comparison between separate cohorts. Five used a prospective design, [81, 83, 89, 95, 98] and four were retrospective.[85, 90, 91, 100] Only one study used propensity score matching to adjust for differences in participant characteristics,[81] and one study performed a comparison between software and cognitive fusion using conditional logistic regression.[98] All other studies reported naïve, unadjusted comparisons.

Six studies compared software fusion alone with cognitive fusion alone[85, 89-91, 95, 98] and two studies reported a comparison between software fusion with concomitant systematic biopsy against cognitive fusion with systematic biopsy.[89, 100] Two studies compared different software fusion technologies against one another; one compared two technologies (Biojet with Urostation),[81] and another compared three (Biojet, Koelis Trinity and Uronav).[83] The following software fusion technologies were evaluated: Biojet (3 studies),[81, 83, 89] Biopsee (one study),[95] BK fusion (two studies),[85, 100] and iSR'obot Mono Lisa (one study).[91] Three studies included a software fusion technology manufactured by KOELIS, including Trinity (one study),[83] Urostation (two studies) [81, 98]

4.4.3.1 Software fusion versus cognitive fusion

Prostate cancer

Five studies compared software fusion with cognitive fusion and reported PCa rates.[85, 90, 91, 95, 98] All three studies that reported a definition of PCa used the same threshold (Gleason score of 6). Their results are presented in Table 12, with further details presented in Appendix 7.

Three studies reported higher test positive rates of PCa for subjects receiving software fusion compared with cognitive fusion; two of those reported that the difference was statistically significant,[91, 98] and one did not report measures of statistical significance.[95] One study found no statistically significant difference between cognitive fusion and software fusion,[85] and one study reported higher test positive rates for cognitive fusion but no measures of statistical significance.[90]

Overall, these five studies broadly agree with the findings of the NMA which showed software fusion was associated with more PCa detection than cognitive fusion. However, the evidence from these five

studies is inconsistent and also at high risk of confounding, notably due to the lack of paired or randomised comparison.

Table 12 Software fusion vs. cognitive fusion, prostate cancer test positive rates (studies not included in the meta-analyses)

	Population	SF technology	Route#	Anaesthesia#	Sample size	PCa definition	Effect estimates	Statistical significance
Delongchamps (2013)[98]	BN	Urostation Touch (Koelis)*	TR	NR	SF: 82 CF: 54	NR	NR&	SF vs. SB: p=0.006 CF vs. SB: p=0.22
Hansen (2018)[95]	BN	Biopsee	TP	GA	SF: 395 CF: 176	NR	SF: 53% CF: 38%	NR
Kaufmann (2018)[91]	BN, RB	Biojet	SF: TP CF: TR	NR	SF: 191 CF: 87	GS: 6	SF: 58.1% CF: 43.7%	p=0.02
Liang (2020)[85]	BN	BK Fusion^	TP	LA	SF: 92 CF: 71	GS: 6	SF: 51.08% CF: 60.56%	p=0.228
Monda (2018)[90]	BN, RB	UroNav	TR	NR	SF: 348 CF: 162	GS: 6	SF: 14.4% CF: 22.8%	NR

For both SF & CF approaches unless otherwise specified; * Also compared to Esaote rigid software fusion system. ^ Predictive Fusion Software. & Probability of detecting cancer undetected by SB against SB as reference was calculated but NR. PCa: prostate cancer; BN: biopsy naïve; SF: Software fusion; CF: cognitive fusion; SB: systematic biopsy; NR: not reported; GS: gleason score

Clinically significant prostate cancer

Five studies compared software fusion with cognitive fusion and reported data on clinically significant PCa test positive rates.[85, 90, 91, 95, 98] All studies defined clinically significant cancer as gleason score of 7 (3+4) or higher. Their results are presented in Table 13, with further detail in Appendix 7.

Two studies reported no statistically significant difference in test positive rates of clinically significant PCa between software fusion and cognitive fusion,[85, 91] whereas one study reported a statistically significant difference in test positive rates favouring software fusion.[98] One study reported a higher rate of clinically significant PCa for cognitive fusion compared with software fusion, although it did not report whether this difference was statistically significant.[91] One study reported similar rates of clinically significant cancers between software fusion and cognitive fusion,[90] and comparable rates of missed, upstaged and equivalent clinically significant biopsy results identified by each targeted biopsy method against concurrent 14-core, systematic biopsy (software fusion, p=0.172).

Although outcomes between these studies are inconsistent and are at high risk of bias overall, they do not show evidence of a significant difference in rates of clinically significant PCa detection between software fusion and cognitive fusion. This evidence is broadly reflective of the meta-analysis findings.

Table 13 Software fusion vs. cognitive fusion, clinically significant prostate cancer biopsy positive rates (studies not included in the meta-analyses)

	Population	SF technology	Route[#]	Anaesthesia[#]	Sample size	CsPCa definition	Effect estimates	p-value
Delongchamps (2013)[98]	BN	Urostation Touch (Koelis)*	TR	NR	SF: 82 CF: 54	GS ≥3+4	NR[&]	SF vs. SB: p=0.001 CF vs. SB: p=0.6
Hansen (2018)[95]	BN	Biopsee	TP	GA	SF: 395 CF: 176	GS ≥3+4	SF: 56% CF: 70%	NR
Kaufmann (2018)[91]	BN, RB	Biojet	SF: TP CF: TR	NR	SF: 191 CF: 87	GS ≥3+4	SF: 80.4% CF: 84.6%	p=0.55
Liang (2020)[85]	BN	BK Fusion^	TP	LA	SF: 92 CF: 71	GS ≥3+4	SF: 35.87% CF: 39.43%	p=0.641
Monda (2018)[90]	BN, RB	UroNav	TR	NR	SF: 162 CF: 348	GS ≥3+4	SF: 27.9%, CF: 27.2%	NR

* Also compared to Esaote rigid software fusion system. [&] Probability of detecting clinically significant cancer undetected by SB against SB as reference was calculated but NR. ^ 'Predictive Fusion Software'CsPCa: clinically significant prostate cancer; BN: biopsy naïve; SF: Software fusion; CF: cognitive fusion; NR: not reported; GS: Gleason score; TR: Transrectal; TP: Transperineal; LA: local anaesthetic

4.4.3.2 Software fusion and systematic biopsy versus cognitive fusion and systematic biopsy

Two studies that were excluded from the meta-analyses compared PCa test positive rates between software fusion with concomitant systematic biopsy, against cognitive fusion with systematic biopsy. [89, 100] Results are summarised in Table 14, with further details in Appendix 7. There was no statistically significant difference in rates of overall PCa and clinically significant cancer detection between the two methods.

Table 14 Software fusion with systematic biopsy vs. cognitive fusion with systematic biospy, prostate cancer and clinically significant prostate cancer test positive rates (studies not included in the meta-analyses)

	Population	SF technology	Route	Anaesthesia	Sample size	Outcome	Effect estimates	Statistical significance
Lockhart (2022)[100]	BN	BK fusion[+]	TP	NR	SF+SB: 97 CF+SB: 186	GS ≥3+4	SF+SB: 53% CF+SB: 66.7%	NR
Stabile (2018)[89]	BN, RB	Biojet	SF: TP or TR CF: TR	NR	SF: 157 CF: 87	PCa (not defined)	SF+SB: 68.2% CF+SB: 58.6%	p=0.2
						CsPCa (GS ≥3+4)	SF+SB: 58% CF+SB: 44.8%	p=0.07

[+] MIM fusion software platform with a BK3000 ultrasound.
PCa: prostate cancer; CsPCa: clinically significant prostate cancer; BN: biopsy naïve; RB: repeat biopsy; SF: Software fusion; CF: cognitive fusion; SB: systematic biopsy; NR: not reported; GS: Gleason score

4.4.3.3 Software fusion versus with software fusion

Two studies that were not included in the meta-analyses compared biopsy test positive rates between software fusion technologies.[81, 83] One study compared Biojet with Koelis Urostation, and one study evaluated three devices: Biojet, Koelis Trinity, and Uronav. Results are summarised in Table 15, with further details in Appendix 7. Both studies found no statistically significant difference in test positive rates of PCa and clinically significant PCa between software fusion devices. Overall, this evidence is consistent with the findings of the meta-analyses.

Table 15 Software fusion vs. software fusion, prostate cancer and clinically significant prostate cancer biopsy test positive rates (studies not included in the meta-analyses)

	Population	SF technology	Route#	Anaesthesia#	Sample size	Outcome	Effect estimates	Statistical significance
Ferriero (2022)[81]	BN, RB	Biojet Urostation	Urostation: TR Biojet: NR	NR	SF: 103 (83)* SF: 211 (83)*	PCa (NR) Per target	SF(Urostation): 69.8% SF (Biojet): 56.6%	p=0.077
						CsPCa (NR) Per target	SF(Urostation): 50.6% SF (Biojet): 50.6%	p=1
Sokolakis (2021)[83]	BN, RB	Biojet Koelis Trinity Uronav	TR	LA	Biojet: 20 Trinity: 20 Uronav: 20	PCa	Biojet: 65% Trinity: 70% Uronav: 65%	p>0.99
						CsPCa (GS ≥3+4)	Biojet: 50% Trinity: 55% UroNav: 50%	p>0.99

For all SF approaches unless otherwise specified; * Values in brackets refer to effective sample sizes following propensity score matching. PCa: prostate cancer; CsPCa: clinically significant prostate cancer; BN: biopsy naïve; RB: repeat biopsy; SF: Software fusion; CF: cognitive fusion; NR: not reported; GS: Gleason score; LA: local anaesthetic.

4.4.3.4 Subgroups

Three subgroups were prespecified in the NICE scope and review protocol: patients with anterior lesions, patients with posterior lesions, and individuals who have had a previous negative prostate biopsy and are referred for a repeat biopsy within 12 months. The review protocol also specified that the following potential factors affecting diagnostic accuracy would be investigated in subgroup analyses: biopsy naïve patients, and operator experience. Test positive rates by PI-RADS groups (3, 4 and 5) were also summarised, although this subgroup was no pre-specified.

NMAs for biopsy-naïve or prior negative biopsy subgroups were not conducted due to the limited number of studies included. Absolute probabilities of being classified as having no cancer or being at different ISUP grades are presented for biopsy-naïve or patients with a previous negative biopsy in Section 4.4.2.1. Whilst it is expected that these characteristics may influence the number of positive cancers detected (due to a different underlying prevalence of cancer in the different populations), there is no evidence that they may affect the relative diagnostic accuracy across biopsy types.

Lesion location

One study[31] reported test positive estimates by lesion location (anterior, posterior), and found no significant differences in test positive rates of PCa and clinically significant PCa between software fusion (Biopsee) and cognitive fusion for posterior and anterior located lesions. The results are summarised in Table 16. Test positive rates were also stratified by other locations (peripheral and transition zones, not reported here) and showed no statistically significant differences between the two methods.

Table 16 Test positive rates of prostate cancer and clinically significant prostate cancer by lesion location (anterior, posterior) in FUTURE

Study	SF technology	Route	Anaesthesia	Lesion location	N of lesions	Outcome (definition)	Test positive rates	Statistical significance
FUTURE (2019)[31]	Biopsee	SF: TP CF: TR	SF: GA CF: LA	Anterior	SF: 37 CF: 25	PCa (NR)	SF: 62.2% CF: 60.0%	p>0.9
						CsPCa (GS: ≥3+4)	SF: 48.6% CF: 44.0%	p=0.6
				Posterior	SF: 35 CF: 46	PCa (NR)	SF: 40.0% CF: 26.1%	p=0.12
						CsPCa (GS: ≥3+4)	SF: 20.0% CF: 26.1%	p=0.7

PCa: prostate cancer; CsPCa: clinically significant prostate cancer; SF: Software fusion; CF: cognitive fusion; GS: Gleason score; NR: Not reported; TR: transrectal; TP: transperineal; GA: general anaesthetic; LA: local anaesthetic.

Repeat biopsy and biopsy naïve patients

Test positive rates for patients receiving a repeat biopsy following a prior negative biopsy and for biopsy naïve patients are presented in Appendix 7, Table 74 and Table 75 respectively. Overall, there was no evidence that software fusion had higher test positive rates compared with cognitive fusion in either subgroup.

Impact of operator experience

One study evaluated how operator experience impacts the cancer biopsy positivity rates.[89] The results are reported in Table 17. Stabile (2018) evaluated the learning curve for the probability of detecting csPCa from three urologists, who each used a different biopsy approach on separate patient cohorts: transrectal cognitive biopsy (operator 1), transrectal software fusion biopsy (operator 2), and transperineal software fusion biopsy (operator 3). Each urologist had performed at least 200 prostate biopsies but were naïve to targeted biopsy techniques. The total number of targeted biopsies performed by operator 1, 2 and 3 were 87, 70 and 87 respectively. Operator experience was defined as the progressive number of targeted biopsies performed by each operator. Stabile (2018)[89] found that there was a sharp increase in the csPCa biopsy positivity rates in the first 60 procedures, where it plateaued, regardless of the biopsy approach. Operator experience was a significant predictor of the

csPCa biopsy positivity rate in targeted cores, which was more pronounced for the operator who conducted transrectal software fusion biopsy compared with the other two biopsy approaches.

Table 17 Impact of operator experience on prostate cancer detection

Study	Pop.	SF technology	Route	Anaesthesia	Sample size	N of targeted biopsies	N cores per ROI	Effect estimates	p-value
Stabile (2018)[89]	BN, RB	Biojet	SF: TP or TR CF: TR	NR	SF: 157 CF: 87	TR SF: 70 TP SF: 87 TR CF: 87	Med (range) SF: 3 (2-3); CF: 2 (2-5)	Learning Curve csPCa detection by operator experience: OR 1.03, 1.06, and 1.01 for operators 1, 2, and 3, respectively csPCa biopsy positivity rate at first procedure to 60th procedure: Operator 1 (TR CF): 30-57% Operator 2 (TR SF): 15-78% Operator 3 (TP-SF): 70-83%	p <0.04

BN, biopsy naïve; RB, repeat biopsy; SF, software fusion; CF, cognitive fusion; ROI, regions of interest; TP, transperineal; TR, transrectal; NR, not reported; csPCa: clinically significant prostate cancer

PI-RADS

Six studies reported test positive rates of PCa stratified by PI-RADS score (3, 4, or 5). All four studies that reported any PCa rates for software fusion and cognitive fusion found no statistically significant differences by PI-RADS score between the two methods.[31, 85, 88, 95] Similarly, the two studies that compared clinically significant rates between software and cognitive fusion subgroups found no difference across PI-RADS subgroups.[31, 95] One study[81] found that test positive rates of any PCa cancer and csPCa were comparable between Koelis Urostation and Biojet after stratifying for PI-RADS score except for PI-RADS Score 4, where the rate of any PCa was higher in the Urostation group compared with Biojet (80% vs 58.1%, respectively for EF and RF groups, p=0.025), and one study[84] found that rates of CsPCa were higher for PI-RADS 4 patients undergoing transperineal biopsy with Biojet compared with transrectal biopsy with Artemis (43.4% vs. 33.3%), but similar for PI-RADS 3 and 5 subgroups. These results are all based on small (n<100) subgroups and may not be reliable.

4.5 Clinical effectiveness results

4.5.1 Biopsy positivity rates

Four studies reported biopsy positivity rates outcomes; [31, 84, 88, 98] their results are presented in Table 18. Three studies compared software fusion with cognitive fusion and one compared different

software fusion biopsies. None of the studies reported what threshold was used to define biopsy positivity rates. Biopsy positivity rates varied widely, from 21.1% to 75% for SF, and from 33.3% to 67% for CF.

Overall, there is no evidence that biopsy positivity rates differ significantly between software fusion and cognitive fusion. Evidence comparing biopsy positivity rates between software fusion devices is inconclusive, as it limited to one study at high risk of confounding.

4.5.1.1 Software Fusion versus Cognitive Fusion

Of the three studies that compared SF with CF, two studies did not find any significant difference in biopsy positivity rates between the two methods;[31, 88] one study found a statistically significant difference in biopsy positivity rates that favoured SF,[98] although its results may be confounded due to the lack of matching or adjustment between the two study arms.

4.5.1.2 Comparisons between software fusion technologies

One study[84] found that the biopsy positivity rate of Biojet was significantly higher than that of Artemis (43.5% vs 21.1% respectively, p= 0.0002). However, this finding is at high risk of confounding, due to differences in biopsy route (transrectal for Artemis, and transperineal for Biojet) and anaesthesia (local for Artemis, and general for Biojet) between the two study arms.

Table 18 Biopsy postivity rates from studies included in the systematic review

Study	Design	Pop.	Biopsy Method			Sample size (participants)	Total n of cores	N cores	N ROI targeted[#]	Biopsy positivity definition	Effect estimates	p-value
			Type	Route	Anaesthesia							
Software fusion vs. cognitive fusion												
Delongchamps (2013)[98]	Consecutive series, between-patient	BN	SF (Urostation Touch)[&] CF	SF & CF: TR	NR	mpMRI +ve Koelis: 82 CF: 54	NR	Med (range) SF: 3 (2-5) CF: 4 (3–10	NR	NR	Median % (IQR) SF: 75% (33-100) CF: 67% (20-86)	p=0.003
FUTURE (2019)[31]	RCT, between patient	RB	SF (Biopsee); CF	SF: TP, CF: TR	NR	157 (SF: 79, CF: 78)	SF: 358 CF: 275	Med (IQR) SF: 4 (3-5) CF: 3 (3-4)	All ROI	NR	Mean % (SD) SF: 31.3% (37.8) CF: 33.3% (42.1)	NR
PAIREDCAP (2019)[88]	Prospective cohort, within patient	BN	SF (Artemis); CF; SB	NR	NR	248	SF: 741 CF: 744	3 cores	Index ROI	NR	SF: 38.1% CF: 33.3% SB: 15.7%*	SF vs CF: NS*
Software fusion vs. software fusion												
Rabah (2021)[84]	RCT, between patient	BN, RB	SF (Artemis), SF (Biojet)	Artemis: TR Biojet: TP	Artemis: LA Biojet: GA	307	Artermis: 403 Biojet: 338	2-4 cores	All ROI	NR	Biojet: 43.5% Artemis: 21.1%	p=0.0002

[#]Unless specified, the number of cores sampled, and number of ROIs targeted related to both targeted biopsy methods. [&] Also compared to Esaote rigid software fusion system.
* The biopsy positivity rate was significantly higher for targeted biopsies (software fusion and cognitive fusion) compared to systematic biopsy p = 0.008
RCT, randomised controlled trial; BN, biopsy naïve; RB, repeat biopsy; SF, software fusion; CF, cognitive fusion; SB: systematic biopsy; ROI, regions of interest; TP, transperineal; TR, transrectal;GA, general anaesthesia; LA, local anaesthesia, IQR: interquartile range; SD: standard deviation; NS: not significant; NR: not reported

4.5.2 Time taken for biopsy procedure

Two studies compared procedure completion rates; both between different SF devices. The results of these studies are presented in Table 19. Procedure completion rates varied widely, from an average of 13 minutes to 41 minutes; this variation is likely due in part to differences in biopsy and anaesthesia methods.

Overall, there is evidence suggesting that duration of biopsy procedures performed transrectally under local anaesthesia, using Biojet or Uronav (rigid registration) is significantly shorter than with Koelis Trinity (elastic registration). However, this finding is based on a single, small study and is not conclusive.

Both of the studies found statistically significant differences in procedure time between SF devices. Sokolakis (2021)[83] found biopsies conducted transrectally under local anaesthesia were significantly faster using Biojet and UroNav devices (both with rigid registration), compared with the Koelis Trinity device (elastic registration). In Rabah (2021)[84], the time taken to conduct the biopsy procedure was significantly shorter using the Artemis device, compared to the Biojet device, although this comparison is at high risk of confounding due to differences in biopsy route and anaethesia method: biopsies conducted with Artemis were performed transrectally under local anaesthesia, whereas biopsies with Biojet were done transperineally under general anaesthesia.

Sokolakis (2021),[83] also compared the time taken to conduct the biopsy procedure by operator experience. Four urologists (two trainees who had completed around 40 TRUS-guided biopsies (junior urologists), and two senior urologists who had completed more than 250 TRUS-guided biopsies; but none had any experience of software fusion) conducted five biopsies with each system. Overall, operative time for the rigid registration fusion devices shorter for the senior urologists compared to the junior urologists, but there were minimal differences in operating time for the elastic registration fusion device.

Table 19. Time taken for biopsy procedure

Study	Design	Pop.	Biopsy Method			Sample size	N cores per ROI[#]	N ROI targeted[#]	Effect estimates	p-value
			Type	Route	Anaesthesia					
Rabah (2021)[84]	RCT, between patient	BN, RB	SF: Artemis, SF: Biojet	Artemis: TR Biojet: TP	Artemis: LA Biojet: GA	Artemis: 165 Biojet: 142	2-4 cores	All ROI	Mean (SD) Biojet: 41.2 mins (±0.7) Artemis: 13 mins (±2.3)	p<0.001
Sokolakis (2021)[83]	Prospective cohort, between patient	BN, RB	SF: Biojet SF: Koelis SF: UroNav	TR (all)	LA (all)	Biojet: 20 Koelis: 20 UroNav: 20	2-3 cores	All ROI	Median (IQR) Biojet: 16 mins (15-18) Koelis: 28 mins (26-29) UroNav: 17 mins (15-20)	p<0.001

[#] Number of cores and number of ROI targeted relate to both targeted biopsy method unless specified otherwise.

BN, biopsy naïve; RB, repeat biopsy; SF, software fusion; CF, cognitive fusion; ROI, regions of interest; TP, transperineal; TR, transrectal;GA, general anaesthesia; LA, local anaesthesia; SD: standard deviation; IQR: interquartile range;

4.5.3 Complications and adverse events

Five studies evaluated the adverse events and complications arising from the prostate biopsy procedure.[31, 83-85, 90] Of those, three studies compared complication rates and adverse events of software fusion and cognitive fusion, and two compared different software fusion devices.

Overall, there is no evidence of a significant difference in safety outcomes between biopsies conducted with software fusion and cognitive fusion, although the evidence is limited by poor reporting and at high risk of confounding due to differences in biopsy routes and anaesthesia methods.

4.5.3.1 Software fusion versus cognitive fusion

Table 20 presents the results of the three studies that compared safety events between software fusion and cognitive fusion. Of those, two found no difference in safety outcomes (severity not reported) between the two fusion methods,[85, 90] and one found a higher rates of grade 1-2 adverse events for patients undergoing cognitive fusion transrectal biopsy under local anaesthesia compared with software fusion transperinal biopsy under spinal/general anaesthesia. As discussed in section 4.3.2, the comparison in this study is at high risk of confounding due to the different biopsy routes and anaesthesia methods.

Table 20 Summary of safety outcomes from studies included in the systematic review

Study	Design	Pop.	Biopsy method			Sample size	N cores per ROI[#]	Total N of cores	Effect estimates			p-value
			Type	Route	Anaesthesia							
FUTURE (2019)[31, 104]	RCT, between patient	RB	SF (Biopsee, TP) vs CF (TR)	SF: TP, CF: TR	SF: GA/ spinal anaesthesia CF: LA	SF: 79 CF: 78	Median (IQR) SF: 4 (3-5) CF: 3 (3-4)	SF: 358 CF: 275	SF vs CF: OR 2.27 [95% CI 1.04-5.00] (Grade 1-2) Grade 1 AEs (SF: 65.8%; CF: 74.4%); Grade 2 AEs (SF: 5.1%; CF: 10.3%) • Grade 1-2 AEs			p <0.05
										SF (%)	CF (%)	
									Haematuria	50.6	74.4	
									Haematospermia	35.4	50.0	
									Rectal bleeding	2.5	5.1	
									UTI	1.3	6.4	
									Fever	2.5	5.1	
									Urinary retention	3.8	5.1	
									Haematoma	3.8	-	
									Lower back pain	1.3	-	

									Atrial fibrillation	1.3	-	
Liang (2020)[85]	Retrospective, between patients	BN	SF (BK Predictive Software) vs CF (Both TP)	TP	LA	SF: 92 CF: 71	4	NR	SF: 2 AEs (1 post-biopsy fever, 1 bacteremia). CF: 2 AEs (2 post-biopsy fever). AE grade not reported. No patients developed severe bleeding, dysuria, vasovagal reactions, or other complications that required to be addressed.			NR
Monda (2018)[90]	Retrospective; before and after study	BN, RB (+ve/-ve)	SF: UroNav vs CF. (Both TR)	NR	NR	SF: 348 CF: 162	NR	NR	% patients with complications: CF: 8.6%; SF: 7.2% AE grade not reported.			p=0.564

[#] Number of cores and number of ROI targeted relate to both targeted biopsy method unless specified otherwise.
RCT, randomised controlled trial; BN, biopsy naïve; RB, repeat biopsy; SF, software fusion; CF, cognitive fusion; ROI, regions of interest; TP, transperineal; TR, transrectal;GA, general anaesthesia; LA, local anaesthesia; AE, adverse event; IQR, interquartile

4.5.3.2 Comparisons between software fusion technologies

Two studies that compared software fusion technologies reported safety outcomes.[83, 84] Both studies found similar rates of AEs. Rabah (2021)[84] found no difference between the rates of urinary retention or haematuria (p = 0.56, p = 0.6 respectively) between two software fusion biopsy devices (Artemis and Biojet), although these results are at high risk of confounding due to differences in biopsy route. Sokolakis (2021)[83] found no severe peri- or post-operative AEs, but mild AEs were reported in most participants, although this was not evaluated statistically.

Table 21 Complication and Adverse Events

Study	**Design**	**Pop.**	**Biopsy method**			**Sample size**	**N cores per ROI[#]**	**Total N of cores**	**Effect estimates**	**p-value**
			Type	**Route**	**Anaesthesia**					
Rabah (2021)[84]	RCT, between patient	BN, RB	SF: Artemis (TR) and Biojet (TP)	Artemis: TR Biojet: TP	Artemis: Local Biojet: General	Artemis:165 Biojet: 142	2-4	Artermis: 403 Biojet: 338	Haematuria: 2 Artemis 1 Biojet Urinary Retention 7 Artemis 8 Biojet Rectile Bleeding 6 Artemis AE grade not reported	p=0.6 p=0.56

Sokolakis (2021)[83]	Prospective cohort, between patient	BN, RB	SF: Biojet vs Koelis vs UroNav (All TR)	TR	Local	Biojet: 20 Koelis: 20 UroNav: 20	2-3	NR	No severe peri- or post-operative AEs. Transient AEs common (hematuria, hematospermia and hematochezia)	NR

[#] Number of cores and number of ROI targeted relate to both targeted biopsy method unless specified otherwise.
BN, biopsy naïve; RB, repeat biopsy; SF, software fusion; CF, cognitive fusion; ROI, regions of interest; TP, transperineal; TR, transrectal; AEs, adverse event; NR, not reported

4.5.4 Operator preferences between software fusion technologies

One study[83] evaluated the usability of software fusion biopsy which found evidence suggesting that rigid systems (Biojet and Uronav) are easier to use compared to the elastic registration system (Koelis) for tranrectal biopsies under local anaesthesia, although this finding is based on a single small study at high risk of bias and is therefore not conclusive.

Sokolakis (2021)[83] compared the impact of operator experience on the usability of three software fusion devices, using a system usability scale: a 100-point scale measuring the learnability and user-friendliness of a given technology, with higher values indicating a device or technology is easier to use.[105] Senior urologists also found that the software fusion devices had better usability compared to the junior urologists. Sokolakis (2021)[83] also compared the usability of the three software fusion devices and found that the rigid systems (Biojet and Uronav) were significantly easier to use compared to the elastic registration system (Koelis).

Table 22 Summary of usabiility

Study	Design	Pop.	Biopsy methods			Sample size	N cores per ROI[#]	Total N of cores	Effect estimates
			Type	Route	Anaes-thesia				
Sokolakis (2021)[83]	Prospective, between patient	BN, RB	SF: Biojet, Koelis, UroNav	TR (all)	LA (all)	Biojet:20 Koelis:20 UroNav:20	2-3	NR	System Usability Scale [Median (IQR)] Total Biojet: 65 (63.8, 68.1); Koelis: 38.8 (37.5,45); UroNav: 72.5 (63.8, 80.6) Junior Urologists Biojet: 65 (65, 65); Koelis: 38.8 (38.1, 39.4); UroNav: 62.5 (61.2, 63.8) Senior urologists Biojet: 68.8 (64.4, 73.1); Koelis: 48.8 (43.1, 54.4); UroNav: 81.2 (80.6, 81.9) p-values NR

[#] Number of cores and number of ROI targeted relate to both targeted biopsy method unless specified otherwise.
BN, biopsy naïve; RB, repeat biopsy; SF, software fusion; CF, cognitive fusion; ROI, regions of interest; TR, transrectal;GA, general anaesthesia; LA, local anaesthesia; IQR, interquartile

4.5.5 Other outcomes

No evidence was found for the following outcomes specified in the protocol: biopsy sample suitability/quality, number of repeat biopsies performed, procedure completion rates, software failure rate, time to diagnosis, length of hospital stay, re-biopsy rate, hospitalization, overall survival, progression free survival, patient- and carer reported outcomes (including tolerability and health-related quality of life), barriers and facilitators to implementations, or cost outcomes.

4.6 Diagnostic accuracy and clinical effectiveness: summary and conclusions

The evidence identified by the systematic review included a total of 3733 patients who received software fusion and 2154 individuals with cognitive fusion from 23 studies. Evidence was included for all devices specified in the protocol, except for Fusion Bx 2.0 and FusionVu. Overall, the evidence for all devices was at high risk of bias. Up to fourteen studies were included in network meta-analyses. Analyses compared the relative diagnostic accuracy of software fusion, cognitive fusion, cognitive fusion with concomitant systematic biopsies, software fusion with systematic biopsies, and systematic biopsies alone.

Our main network meta-analyses looked at how cognitive fusion compares to software fusion in classifying patients across the range of ISUP grades. Results must be cautiously interpreted due to the high risk of bias, but suggest that patients undergoing cognitive biopsy may show: i) a higher probability of being classified as not having cancer, ii) similar probability of being classified as having non-clinically significant cancer (ISUP grade 1), and iii) lower probability of being classified at higher ISUP grades, particularly ISUP 2. Similar results were obtained when comparing between same biopsy methods where both were combined with systematic biopsy.

Additional meta-analyses of cancer detection rates suggest that, compared with cognitive fusion biopsy, software fusion may identify more prostate cancer (any grade) (OR 1.30; 95% CrI 1.06, 1.61) and more non-clinically significant cancer (ISUP 1) (OR 1.98; 95% CrI 1.28, 3.06). Adding systematic biopsy to cognitive or software fusion may increase the detection of all prostate cancer and of clinically significant cancer, and from this evidence there is no suggestion that software fusion with concomitant systematic biopsy is superior to cognitive fusion with systematic biopsy.

Meta-analyses by individual device showed that compared with cognitive fusion biopsy, Biojet are Urostation are associated with a higher detection of prostate cancer overall, and that Biojet is associated with more clinically significant cancer, although only one study of Biojet was included in the meta-analyses. The evidence for all other software devices was insufficient to evaluate their accuracy compared with cognitive fusion reliably, or to assess whether some software fusion technologies are more accurate than others.

Compared to cognitive fusion biopsy, the meta-analyses showed evidence that systematic biopsy may detect more cancers overall than cognitive biopsy alone, but similar rates of clinically significant cancer overall. There was large uncertainty in all estimates due to the limited evidence, particularly for higher ISUP grades and by individual device. Results from studies excluded from the meta-analyses broadly reflected these findings. Compared with cognitive fusion, there was no evidence that the accuracy of software fusion may differ by lesion location, or between biopsy naïve and prior negative biopsy patients, or according to operator experience.

The applicability of the evidence for KOELIS Trinity is uncertain, as it was almost entirely informed by evaluations of a previous version (KOELIS Urostation) without integrated ultrasound. The applicability of the evidence for Biopsee is also limited due to the lack of evaluations under local anaesthesia. There is no evidence comparing the accuracy of Fusion Bx 2.0 and FusionVu with cognitive fusion, and no evidence for these devices were eligible for inclusion in the indirect comparisons.

Evidence for all other protocol specified outcomes was limited and inconclusive. Overall, there is no evidence that biopsy positivity rates differ significantly between software fusion and cognitive fusion, or between software fusion devices. There was some evidence that systems with rigid registration (Biojet or Uronav) are easier and significantly faster to use than elastic registration (KOELIS Trinity), although this is informed by a single, small study and is not conclusive. Overall, there is no evidence of a significant difference in safety outcomes between biopsies conducted with software fusion and cognitive fusion or between software fusion devices, although the evidence is limited by poor reporting and at high risk of confounding.

No relevant evidence was found for the following outcomes: biopsy sample suitability/quality, number of repeat biopsies performed, procedure completion rates, software failure rate, time to diagnosis, length of hospital stay, time taken for MR image preparation, subsequent prostate cancer management, re-biopsy rate, hospitalization, overall survival, progression free survival, patient- and carer reported outcomes (including tolerability and health-related quality of life), barriers and facilitators to implementations.

4.7 Additional evidence to inform model structure and parameterisation

Additional evidence was required to inform a number of economic parameters, including 1) prostate cancer prevalence; 2) distribution of test results for cognitive and software fusion broken down by Gleason grade; 3) test accuracy of cognitive and software fusion; 4) long-term evidence on outcomes from management strategies for patients with prostate cancer. In addition to the systematic review of diagnostic accuracy and clinical effectiveness, targeted reviews were conducted to identify the most

relevant evidence to inform these parameters. This section describes these reviews and the evidence identified.

4.7.1 Review of additional prevalence, test results and diagnostic accuracy evidence

Studies included in the systematic review of diagnostic accuracy and clinical effectiveness were reviewed to identify suitable evidence to inform the following economic model parameters: 1) PCa prevalence, estimated from a 'gold-standard' test (template mapping or saturation biopsy with at least 20 cores) and with sufficient granularity (by Gleason group); 2) distribution of test results for cognitive or software fusion MRI in PI-RADS 3+ by Gleason group; and 3) accuracy of cognitive or software fusion MRI in PI-RADS 3+ patients against a 'gold-standard' test, i.e. comparative studies against template mapping or saturation biopsy for which a composite endpoint could be derived from the results of both tests.

Due to the lack of evidence from 'gold-standard' tests identified in the systematic review, additional targeted, pragmatic searches were conducted. References from a recent Cochrane systematic review, which included studies on the diagnostic accuracy of MRI-targeted biopsy against template-guided biopsy, were checked for further evidence.[66] As the searches in the Drost (2019) [66] review were limited to July 2018, pragmatic searches of PubMed and Google scholar were conducted to identify more recent studies. This search included the following search terms: ((((template mapping) OR (saturation)) AND (biopsy)) AND (prostate)) AND (fusion biopsy).

Ten studies were considered potentially eligible to inform at least one of the model parameters of interest. Their characteristics are summarised in Table 23. Studies were prioritised according to the applicability of their population to the NHS. The following sections provide further details on the prioritisation and limitations of studies informing each of the three model parameters.

Table 23. Study characteristics from targeted review of prevalence, distribution of test results and test accuracy

Study	Study Design	Sample Size	Population	Biopsy Test 1	Biopsy Test 2
Mannaerts (2020)[106]	Prospective, Within patient	142	Naïve	SF (Artemis)	SB
PAIREDCAP (2019)[88]	Prospective, Within patient	248	Naïve	SF (Artemis)	CF
Izadpanahi (2021) [82]	RCT, Between patient	199	Naive	SF (Artemis) + SB	CF + SB
Filson (2016)[96]	Prospective, Within-patient	538#** (273 naïve)	Naïve, repeat or active surveillance**	SF (Artemis)	SB
Alberts (2018)[80]	Prospective, Within-patient	48	Naïve, repeat**	SF (UroStation)	SB
Mortezavi (2018)[107]	Retrospective	291*	Naïve, repeat or active surveillance	SF (BiopSee)	TTMB
Zhou (2018)[108]	Prospective, Between patient	153	NR	SF (Hitachi) or CF	TTMB

Simmons (2018)[109]	Prospective, Within patient	200	Repeat	SF (SmartTarget) or CF	TTMB
Hansen (2016)[110]	Retrospective	289#*	Naïve, repeat or active surveillance**	SF (BiopSee)	TTMB
Kesch (2017)[111]	Prospective, Within-patient	172	Naïve, repeat or active surveillance	MRI targeted (both software and cognitive)	TTMB

* MRI +ve lesion (PI-RADS ≥ 3)
# Excluding patients currently on active surveillance
**despite included in the study, results by previous biopsy experience were separable.
SF: software fusion; CF: cognitive fusion; SB: systematic biopy; TTMB: template-guided mapping biopsy

4.7.1.1 Prevalence

We were unable to identify any population-level evidence on the prevalence of PCa by Gleason group. Of ten studies identified in our targeted review, five studies compared MRI targeted biopsy compared to a template-guided biopsy (template mapping or saturation biopsy (Table 24). [107-109, 111], Hansen, 2016 #2158} The template-guided biopsy does not present perfect accuracy, as the test's accuracy depends on the intensity of cores taken and core (see section 4.1.2.4). Therefore, to approximate prevalence, and given the assumption of negligible false positive results to biopsy, we used a 'composite' reference standard combining the template-guided biopsy with the other biopsy method investigated in each study. The results from the five studies included are shown in Table 24.

Table 24. Prevalence estimates based on Gleason Grade from studies identified in targeted review, using a composite reference standard (PI-RADS ≥3)

	Hansen (2016)[110] Proportion (N)			Zhou (2018)[108]	Simmons (2018)[109]	Mortezavi (2018)[107]	Kesch (2017)[111]
GG	**BN**	**RB**	**AS**	**Proportion (N)**	**Proportion (N)**	**Proportion (N)**	**Proportion (N)**
0	0.306 (26)	0.461 (94)	0.101 (9)	0.340 (52)	0.075 (15)	0.237 (69)	0.276 (35)
1	0.235 (20)	0.181 (37)	0.393 (35)	0.163 (25)	0.210 (42)	0.12 (35)	0.173 (22)
2	0.212 (18)	0.191 (39)	0.270 (24)	0.190 (29)	0.675 (135)	0.285 (83)	0.378 (48)
3	0.129 (11)	0.083 (17)	0.157 (14)	0.131 (20)		0.155 (45)	0.079 (10)
4 or 5	0.118 (10)	0.083 (17)	0.079 (7)	0.176 (27)	0.04 (8)	0.203 (59)	0.094 (12)
N	85	204	89	153	200	291	127

GG: Gleason grade; BN biopsy naïve; RB: repeat biopsy following prior negative biopsy; AS: active surveillance

The results show considerable variation between studies, with for example, the prevalence of no cancer varying between 7.5% and 34% across studies and the prevalence of GG 4 or 5 cancer from 4% to 20%. The reasons for this heterogeneity are unclear, and may arise from the significant clinical diversity across studies, including in participants (settings of care), diagnostic tests (and in the protocols for their implementation) and outcomes, and/or from the methodological diversity across studies, including variability in study design and risk of bias. The results from Hansen (2016)[110] suggest that the position of patients in the pathway may be a significant source of heterogeneity.

4.7.1.2 Distribution of test results obtained with cognitive fusion or software fusion biopsy

The ten studies identified in the targeted review were potentially relevant to inform the distribution of test results obtained with cognitive fusion or software fusion biopsy. Four studies were initially excluded because their population was not considered representative of the NHS.[80, , 107, , 109, , 111] Mortezavi (2018) [107] and Kesch (2017) [111] included patients under active surveillance. Simmons (2018) [109] only included patients with a repeat biopsy, and patients in Alberts (2018) [80] were selected from a population-wide screening programme.

The distribution of test results obtained from a targeted biopsy for the remaining five studies are presented in Table 25. There is considerable heterogeneity in the proportion of patients identified in each Gleason Grade group across the studies.

Table 25. Distribution of test results for potentially included studies identified from the targeted review (Biopsy naïve population only). Results presented as proportion (N)

	Mannaerts [106]	PAIREDCAP[88]		Izadpanahi[82]		Zhou[108]	Hansen (RB) [110]	Filson (BN)[96]		Filson (RB)[96]	
GG	SF (Artemis)	CF	SF (Artemis)	CF + SB	SF + SB	Mixed CF/SF	SF (BiopSee)	SF (Artemis)	SF + SB	SF	SF+SB
0	0.140 (7)	0.379 (94)	0.286 (71)	0.690 (69)	0.556 (55)	0.503 (77)	0.642 (131)	0.469 (128)	0.355 (97)	0.687 (182)	0.585 (155)
1	0.040 (2)	0.153 (38)	0.173 (43)	0.190 (19)	0.253 (25)	0.105 (16)	0.103 (21)	0.165 (45)	0.220 (60)	0.087 (23)	0.151 (40)
2	0.380 (19)	0.21 (52)	0.282 (70)	0.060 (6)	0.131 (13)	0.118 (18)	0.167 (34)	0.198 (54)	0.223 (61)	0.102 (27)	0.117 (31)
3	0.280 (14)	0.157 (39)	0.161 (40)	0.050 (5)	0.03 (3)	0.111 (17)	0.088 (18)	0.168 (46)	0.201 (55)	0.125 (33)	0.147 (39)
4/5	0.160 (8)	0.101 (25)	0.097 (24)	0.010 (1)	0.03 (3)	0.163 (25)	0.118 (10)				
N	50	248	248	100	99	153	204	273	273	265	265

SF: Software fusion; CF: cognitive fusion; SB: systematic biopsy; BN: biopsy naïve; RB: repeat biopsy

Therefore, to ensure that the distribution of Gleason Grades is representative to the NHS population, the remaining six studies' eligibility criteria were compared to determine which was most representative to NHS practice. According to the NICE guideline NG131 and the PCa diagnostic pathway, patients are referred if their prostate-specific antigen levels are above the age-specific reference range (which, for men aged 50-69 is a PSA level of >3.0ng/ml) or if their prostate feels malignant (hard, or lumpy) on digital rectal examination.[10, 17] Furthermore, this DAR is focused on patients with mpMRI visible lesions (PI-RADS 3+), who are biopsy naïve, or are undergoing a repeat biopsy (after a negative result). Table 26 summarises the study eligibility criteria and participant characteristics, and the decisions for inclusion/exclusion.

Table 26. Population eligibility criteria for studies that are considered for use to inform the distribution of test results

Study	Country	PSA-level		DRE Exam	Naïve/Repeat	Considerations for inclusion or exclusion
		Eligibility Criteria	Included Patients			
Mannaerts (2020)[106]	N'lands	≥3.0 – 20ng/mL	Median (IQR) 6.2 (4.7-8.0)	Suspicious DRE	100% BN	Similar referral pathway: appropriate PSA cut-off & suspicious DRE
PAIREDCAP (2019)[88]	USA	<25ng/mL	Median (IQR) 6.2 (4.6-8.2)	Suspicious DRE	100% BN	Similar referral pathway: appropriate PSA cut-off & suspicious DRE
Filson (2016)[96]	USA	'Elevated PSA'	Median (IQR) Naïve: 5.8 (4.4-8.1) Repeat: 7.6 (5.0-11.5)	Suspicious DRE	33% BN 32% RB 35% AS	Unclear PSA cut-off, but similar PSA of included patients. Less granularity in Gleason Grades (only data on Grade 3+).
Hansen (2016)[110]	UK	'Elevated PSA'	Median (IQR) Naïve: 6.2 (4.8–8.6) Repeat: 7.8 (4.8-8.6)	Suspicious DRE	20% BN 55% Repeat 25% AS	Unclear PSA cut-off, but similar PSA of included patients. Greater proportion of patients with repeat biopsy, and number of naïve patients is small.
Zhou (2018)[108]	China	> 4ng/mL	Median (IQR) 9.5 (6.5–15.5)	Suspicious DRE	100% BN	Similar referral pathway: appropriate PSA cut-off & suspicious DRE. Concerns regarding high baseline PSA levels in the included patients. Differences in healthcare systems between UK and China.
Izadpanahi (2019)[82]	Iran	>2-10ng/dL	Mean (SD) 6.1 ng/dL (1.3)	Suspicious DRE	100% BN	Concerns regarding reporting of PSA-levels (report ng/dL). Limiting PSA levels to <10ng/dL was not deemed representative of UK practice. Differences in healthcare systems between UK and Iran.

PSA, prostate specific antigen; DRE, digital rectal examination; IQR, interquartile range; SD, standard deviation; BN, biopsy naïve; AS, active surveillance

Two studies[82, 108] were not deemed to be appropriate for use in this analysis. Izadpanahi (2019)[82] limited their population to patients with a PSA > 10ng/mL; and the population in Zhou (2018)[108] had a considerably higher baseline PSA compared to the other studies. In addition, the settings of these studies (Iran and China) may not be reflective of NHS practice.

The remaining four studies[88, 96, 106, 112] were deemed to be most similar to NHS practice, based on population eligibility criteria. All studies applied focused on patients with an elevated PSA and included patients who were referred for suspicious DRE. We considered that only biopsy naïve patients should be included in the analysis, as the vast majority (~90%) of patients in NHS practice will be receiving a first biopsy. Therefore, in the studies where separable data were available, we only included the biopsy naïve patients, as the proportion of patients with repeat biopsy was often high.

4.7.1.3 Accuracy of cognitive fusion or software fusion biopsy

In order to determine the accuracy of cognitive fusion or software fusion biopsy, studies which compared MRI-targeted biopsy (software fusion and/or cognitive fusion) against template or

saturation biopsy were identified. To determine true disease status as closely as possible, patients were reclassified according to a composite reference standard from both tests. Out of four studies,[107-110] two provided test accuracy data with the required granularity by Gleason Group.[107, 108] The characteristics of these two studies are summarised in Table 23. Zhou (2018) [108] compared software fusion biopsies (including both software fusion biopsy [29% of patients] and cognitive fusion biopsy [71% of patients]) with template-guided transperineal prostate saturation biopsy, although the study did not provide accuracy data for software fusion biopsy and cognitive fusion biopsy separately. Mortezavi (2018), on the other hand, does provide data on accuracy specifically for software fusion biopsy compared to transperineal template saturation prostate biopsy. Mortezavi (2018)[107] includes patients who are on active surveillance, who are likely to have a different Gleason score distribution compared with biopsy naïve and prior negative biopsy patients. However, as accuracy evidence is conditional on true disease status, any such differences in the patient population included are not likely to have a significant impact on conditional accuracy estimates.

Table 27 provides the computed conditional (accuracy) probabilities of patients being identified at a particular grade with MRI-fusion given a particular true disease status given by the composite TMB and MRI-fusion results.

Table 27. Accuracy probability for software fusion and/or cognitive fusion against composite template mapping biopsy and MRI-fusion results. Results are presented as N (%).

	Zhou (2018)[108] (Software and cognitive fusion)					Mortezavi (2018)[107]				
Composite reference standard	No cancer	GG1	GG2	GG3	GG 4 or 5	No cancer	GG1	GG2	GG3	GG 4 or 5
No cancer	52/52 (1)	0	0	0	0	69/69 (1)	0	0	0	0
GG1	11/25 (0.44)	14/25 (0.56)	0	0	0	24/35 (0.69)	11/35 (0.31)	0	0	0
GG2	13/29 (0.45)	1/29 (0.03)	15/29 (0.52)	0	0	21/83 (0.25)	17/83 (0.20)	45/83 (0.54)	0	0
GG3	1/20 (0.05)	1/20 (0.05)	2/20 (0.10)	16/20 (0.80)	0	10/45 (0.22)	2/45 (0.04)	9/45 (0.20)	24/45 (0.53)	0
GG 4 or 5	0/27 (0)	0/27 (0)	1/27 (0.04)	1/27 (0.04)	25/27 (0.93)	6/59 (0.1)	2/59 (0.03)	7/59 (0.12)	8/59 (0.14)	36/59 (0.61)

GG, Gleason Grade

The results show significant heterogeneity, with Zhou identifying a higher accuracy at GG 3 and above. To aid interpretation of these results, we next describe the two further studies which are UK based and therefore have higher representativeness than both Zhou (2018) or Mortezavi (2018).

Two UK studies – Simmons (2018)[109] and Hansen (2016)[110] – did not report results with the necessary disaggregation of Gleason Grade. Simmons (2018)[109] reports a within-patient comparison (secondary

analysis of PICTURE trial), of TMP biopsy and targeted biopsy (mixture of cognitive and software fusion) but only reported Gleason Grade 1, 2-3 and 4-5 (reported in Table 28). The full accuracy matrix by Gleason group could not be retrieved for Hansen (2016) [110], but sensitivity at Gleason Group thresholds of 1 or above, 2 or above and 3 or above could be calculated, against a composite reference standard. The table below compares these sensitivity values, with the results from the studies for which fuller reporting of the accuracy matrices was available (Table 29). As Table 29 shows, there is also some variation in the sensitivity results between the UK studies. The results from Mortezavi (2018)[107], are more similar to Hansen (2016)[110] at Gleason grade 3 or above, whereas the results of Zhou (2018)[108] are more similar to Simmons (2018)[109] at the lower grade groups. It is therefore unclear what are the relevant source(s) for the between study heterogeneity observed between Zhou[108] and Mortezavi[107], and the representativeness of both these studies to the UK context is uncertain.

Table 28. Accuracy probability for MRI-fusion against composite template mapping biopsy and MRI-fusion results – Simmons (2018)[109]

	MRI fusion biopsy (mix of software and cognitive)			
Composite reference standard	No cancer	GG1	GG 2 or 3	GG 4 or 5
No cancer	15/15 (1)	0	0	0
GG1	25/42 (0.60)	17/42 (0.40)	0	0
GG2 or 3	15/135 (0.11)	21/135 (0.16)	99/135 (0.73)	0
GG 4 or 5	1/8 (0.13)	0/8 (0.00)	2/8 (0.25)	5/8 (0.63)

Table 29. Sensitivity of MRI-fusion biopsy against reference standard for UK studies, compared to studies with fuller reporting of accuracy matrices.

	Sensitivity against composite reference standard			
	Hansen 2016[110]	**Simmons**[109]	**Mortezavi**[107]	**Zhou**[108]
GG>=1	0.670 (73/109)	0.778 (144/185)	0.725 (161/222)	0.752 (76/101)
GG>=2	0.712 (52/73)	0.741 (106/143)	0.690 (129/187)	0.789 (60/76)
GG>=3	0.529 (18/34)	NA	0.654 (68/104)	0.894 (42/47)

4.7.2 Review of long-term evidence

To inform economic model parameters on morbidity and mortality outcomes for prostate cancer patients, a targeted, pragmatic review was conducted. Searches included reference checking of evidence reviews informing NICE guidance on the management of prostate cancer (NG131), [10] references included in the PROMIS economic analysis, targeted searches for relevant Cochrane reviews in CDSR, and citation searches to identify the most up-to-date follow-up data. Studies evaluating long-terms survival and disease progression outcomes in prostate cancer patients according to prognosis status`, either under active surveillance or receiving radical treatment recommended by

NICE[12] and described in section 2.2.4 were included. Priority was given to larger randomised controlled trials with at least two years of follow-up, individual patient data (IPD) meta-analyses, and large UK cohort studies. Fourteen studies, including 12 RCTs,[55, 59-61, 113-120] one IPD meta-analysis [121] and one cohort study[122] were identified and are listed in Appendix 8. Table 30 provides a brief summary of key trials considered most reflective of current NHS practice. The process for prioritising the final set of studies included in the model is described in Section 5.

Three RCTs evaluated the effect of radical prostatectomy in relation to an observation-based strategy in clinically localised prostate cancer: SPCG4, PIVOT and ProtecT.[55, 119, 120] The comparators differed across trials between observation (PIVOT), watchful waiting (SPCG4) and active monitoring (ProtecT). SPCG4 and PROTecT both included patients with localised, non-metastatic cancer, and PIVOT included low-to-high risk prostate cancer patients. PROTecT was conducted in the UK, PIVOT in the USA and SPCG4 in Sweden, Finland, and Iceland. Follow-up duration ranged from 10 years (PROTecT) to 29 years (SPCG4). PROTecT was the most recent study (1999 to 2009, compared with 1994-2002 for PIVOT and 1989-1999 for SPCG4). None of the studies used mpMRI to diagnose patients.

Only SPCG4 found a significant effect for prostatectomy on overall survival, with the more contemporary studies not identifying an effect on all-cause mortality. ProtectT, which compared radical prostatectomy, radiotherapy and a passive management strategy (active monitoring) found that despite surgery and radiotherapy being associated with lower incidences of disease progression and metastases than active monitoring, at a median of 10 years, prostate-cancer–specific mortality was low irrespective of the treatment assigned, with no significant difference among treatments.

Of the trials identified that focused on treatments for intermediate to high risk disease, STAMPEDE was the largest, most recent and only study conducted in the UK.[59] STAMPEDE evaluated treatments for high risk or metastatic or recurrent cancer.[123] A large-UK based RCT of 2962 men conducted between 2005 and 2013 with a median follow-up of 6.5 years, evaluated three drug treatment combinations for high risk or metastatic cancer including zoledronic acid and docetaxel, as used in addition to standard of care. While zoledronic acid showed no evidence of survival improvement, docetaxel led to improved survival and an increase in adverse events.

Other trials with high-risk and/or metastatic disease included HYPO-RT-PC, GETUG-12 and TAX3501.[60, 61, 117] HYPO-RT-PC compared hypofractionated radiotherapy with conventional radiotherapy in 1180 intermediate to high-risk cancer patients and found that hypofractionated radiotherapy was non-inferior in terms of failure-free survival. GETUG-12 and TAX3501. GETUG-12 evaluated the effectiveness of adding docetaxel, zoledronic acid/estramustine, or both to first-line long-term hormone therapy in patients with high risk prostate cancer (Gleason 8-10) and TAX-3501

evaluated the addition of docetaxel to leuprolide against leuprolide alone in metastatic patients following radical prostatectomy. Both were smaller (GETUG-12 included 413 participants, and TAX-3501 had 228 participants). At median follow-up of 12 years, GETUG-12 found that docetaxel-chemotherapy reduces the risk of clinical relapse or mortality in high risk prostate cancer. TAX-3501 was terminated at 3.4 years and was underpowered to detect differences in PFS between study arms.

Additionally, evidence was sought on UK studies reporting outcomes by the five-stage Cambridge Prognostic Group (CPG) risk stratification system that currently supports treatment decisions in the UK NHS.[12] Only one large cohort study was identified.[122] The study included diagnostic data from 10,139 men with non-metastatic prostate cancer from the Public Health England National Cancer Registration Service and had a median follow-up of 6.9 years, and found that a five-stratum risk stratification system outperformed the previous three-stratum risk stratification system used in the UK in predicting the risk of prostate cancer death at diagnosis in men with primary non-metastatic prostate cancer.

Overall, there is relevant evidence on the effectiveness of radical vs. 'conservative' treatment options in delaying progression to metastatic disease, despite the limited observed impacts on mortality. The most contemporary and relevant evidence is from ProtecT, a recent, UK based study.[124] Although there is UK based evidence favouring the prognostic ability of a 5-level score for prostate cancer mortality, there is no evidence on treatment effectiveness stratified by CPG scores.

Table 30 Summary of potentially eligible long-term evidence for prostate cancer considered to parametrise the economic model

Study	Patient group, enrolment period	Location	Design, interventions if RCT	Mortality and disease progression related outcomes, maximum FU	Conclusions
Bill-Axelson(2011) [SPCG4][120]	localized disease [1989 -1999]	Sweden, Finland, Iceland	RCT, watchful waiting vs. radical prostatectomy	Reported at 15 years: -all cause mortality -PC death. -Distant metastases -local progression	Radical prostatectomy was associated with a reduction in the rate of death from prostate cancer
Wilt (2012) [PIVOT][119]	localized disease [1994 -2002]	USA	RCT, -Observation vs. radical prostatectomy	Reported at 10 years: - all cause mortality -PC death -bone metastases	Prostatectomy did not significantly reduce all-cause or prostate-cancer mortality
James (2015) [STAMPEDE] [123]	Metastatic disease [2005 and 2014]	UK and Switzerland	RCT, SOC arm (androgen deprivation therapy)	Reported at 5 years: -Failure-free survival - all cause mortality	Survival remains disappointing in men presenting with M1 disease who are started on only long-term androgen deprivation therapy

James (2016) [STAMPEDE][59]	high-risk and metastatic disease [2005 and 2013]	As above	RCT, -SOC as above vs. SOC +xoledronic acid vs. SOC + docetaxel, vs. SOC + zoledronic acid and docetaxel	Reported at 7 years: -Overall survival -failure-free survival	Zoledronic acid showed no evidence of survival improvement Docetaxel showed evidence of improved survival accompanied by an increase in adverse events.
Hamdy (2016) [ProtecT][55]	localized disease [1999-2009]	UK	RCT, active monitoring vs. radical prostatectomy, vs. radiotherapy	Reported at 10 years: -Adherence -PC death -All cause mortality -Metastases -disease progression	No significant difference among -active monitoring -radical prostatectomy -radiotherapy
Bryant (2020) [ProtecT][124]	As above	As above	As above	Reported at 10 years -disease progression	There are differences in risk categorisation between men who progressed during ProtecT and those that did not. Different grade, low/intermediate/high risk.
Widmark (2019) [HYPO-RT-PC] [117]	intermediate-to-high-risk prostate cancer [2005 – 2015]	Sweden and Denmark	-Ultra-hypofractionated vs. conventionally fractionated radiotherapy	Reported at 5 years: -failure-free survival - disease-free survival, - PC survival, -overall survival,	Ultra-hypofractionated radiotherapy is non-inferior to conventionally fractionated radiotherapy
Gnanapragasam (2016)[122]	localized or locally advanced disease [2000 – 2010]	UK	Observational, exploring the prognostic ability of 5 levels of CPG scores	Reported at 13.7 years: -PC death Reported at 9.6 years: -all cause mortality	The five-stratum CPG system outperforms the standard three stratum risk system in predicting the risk of PC death

RCT, randomised controlled trial; PC, prostaste cancer; SOC, standard of care; CPG, Cambridge prognostic group

5 ASSESSMENT OF EVIDENCE ON THE COST EFFECTIVENESS OF SOFTWARE FUSION BIOPSY

5.1 Overview

In the next sections, we provide an overview of published cost-effectiveness studies on the use of software fusion biopsy systems in comparison with cognitive fusion for targeted prostate biopsy (Sections 5.2 and 5.5), to determine generalisability of the evidence to inform this assessment's decision problem. In addition, this section presents a targeted review of diagnostic cost-effectiveness studies (Sections 5.3 and 5.6), which model prostate biopsy procedures to identify prostate cancer (same point in the diagnostic pathway as the interventions in this assessment). This targeted review is done with the aim to support the conceptualisation and parameterisation of a *de novo* decision analytic model.

5.2 Methodology of the cost-effectiveness of software fusion biopsy for suspected prostate cancer

The methodology of the systematic review of published cost-effectiveness studies comparing software fusion biopsy systems with cognitive fusion for targeted prostate biopsy in men with suspected prostate cancer is described below. The review aimed to assess the generalisability of existing evidence to the decision problem defined by the NICE DAR scope, and provide a brief overview of the model structure, parameterisation, and results. Titles identified for inclusion in this review, are subsequently included in the review to inform the conceptualisation and development of the *de novo* model alongside other studies.

5.2.1.1 Literature searches

The results of the systematic literature searches carried out to inform the clinical effectiveness of technologies described in the Section 4.1 were used to identify relevant cost-effectiveness studies of software fusion systems compared to cognitive fusion for targeted biopsy in men with suspected prostate cancer.

5.2.1.2 Study selection

Full economic evaluations that consider both costs and consequences (including cost-effectiveness, cost-utility and cost-benefit analyses) were considered for inclusion. A broad range of economic evidence on the use of MRI fusion systems was considered eligible, including economic evaluations conducted alongside trials, studies using modelling approaches and analyses of administrative databased. The inclusion criteria also defined the:

- Population as men with an elevated PSA level and/ or abnormal DRE who had suspicious lesion(s) detected by (bi or multiparametric) MRI;
- Interventions as targeted transperineal or transrectal prostate biopsy using MRI fusion software with or without systematic biopsy, under local or general anaesthesia intervention;
- Comparators as targeted transperineal or transrectal prostate biopsy using cognitive fusion with or without systematic biopsy, under local or general anaesthesia.

Studies reporting only resource use, costs or health-related quality of life were excluded from the review, but considered to support the parametrisation of the *de novo* model.

The information submitted by the companies in response to NICE and the EAG's requests for information was also reviewed to identify economic studies that complied with the inclusion criteria described above.

Studies identified by the search strategies (see Appendix 1) were screened and selected through a two-stage process: i) titles and abstracts identified by the bibliographic search were screened for possible inclusion, and ii) full texts of potentially relevant records were obtained and screened for inclusion. The process was performed independently by two researchers (HP and AD) with any disagreement resolved by consensus.

5.2.1.3 Quality appraisal

Cost-effectiveness evidence selected for inclusion was quality assessed using a checklist tool developed for the assessment of model-based economic evaluations of diagnostic tests.[125]

5.2.1.4 Synthesis of evidence

The characteristics and key findings of the included economic evidence were narratively summarised and tabulated for comparison. The extracted information included:

- The perspective of analysis;
- The comparators and its positioning in the diagnostic pathway, study population and setting, main analytic approaches (e.g., analysis of individual patient data/ decision-analytic model), primary outcomes of the economic analysis;
- Details of adjustment for health-related quality of life (HRQoL), resource usage (direct and indirect costs);
- Estimates of incremental cost-effectiveness and how uncertainty was quantified (e.g., deterministic/ probabilistic sensitivity analysis).

The relevance of existing economic evidence to the current decision problem in the NICE DAR scope was assessed based on:

i. Consistency with the decision problem being considered in this assessment, including relevance to the UK;
ii. Relevance of outputs for decision making (i.e., to estimate long-term NHS costs and QALYs based on morbidity and mortality associated with prostate cancer tailoring according to patient prognosis and preferences); and
iii. The model flexibility which allows the consideration of different subgroups (e.g., patients with previous negative biopsy results) and potential effect modifiers of diagnostic accuracy (e.g., operator experience).

5.3 Methodology of the additional targeted reviews to support model conceptualisation

Given an expected dearth of evidence on the cost-effectiveness of biopsies using software fusion biopsy systems compared to biopsies using cognitive fusion in the UK context, we performed

additional targeted reviews of cost-effectiveness evidence of diagnostic strategies at the point of biopsy to support the model conceptualisation. These aimed to i) identify value components of the biopsy approaches, ii) characterise alternative mechanisms of evidence linkage from disease prevalence, diagnostic accuracy, choice of treatment to final outcomes, and iii) identify any UK relevant sources of evidence.

5.3.1.1 Literature searches

We screened cost-effectiveness modelling studies identified by the main search described in Section 4.1 to identify evaluations of diagnostic strategies in the same diagnostic pathway position proposed for software fusion biopsy systems (i.e., at the point of biopsy), but which do not fulfil the full inclusion criteria for the population, interventions and comparators defined for the main cost-effectiveness review (Section 5.2). We also considered for inclusion cost-effectiveness modelling studies identified in the cost-effectiveness reviews conducted for a previous assessment of the cost-effectiveness of transperineal biopsy for diagnosing prostate cancer recently developed to inform NICE guidance.[126] Studies included in the review of cost-effectiveness studies in scope with this assessment (see Section 5.2) were also included in the targeted review.

5.3.1.2 Study selection

We included studies considered potentially informative for the model conceptualisation and for the identification of relevant input sources of evidence with a particular emphasis on those used in UK based or UK generalisable models. The relevance of these studies to inform the model conceptualisation under the current decision problem was assessed as described in Section 5.5.

5.3.1.3 Quality appraisal

Given the pragmatic nature of this review and its aims, identified studies did not undergo a formal quality appraisal.

5.3.1.4 Synthesis of evidence

The studies identified as potentially relevant were summarised in tabular form. A subset of the studies identified was selected for detailed extraction, if they were model-based cost-effectiveness studies which complied with at least the following criteria:

- UK relevant evaluations of alternative prostate biopsy approaches;
- UK policy relevant assessments of diagnostic tests for prostate cancer;
- or evaluations comparing alternative MRI-influenced biopsy approaches.

The value of diagnostic technologies is to a large extent dependent on how downstream clinical management choices based on diagnostic information impact on final outcomes. Therefore, most of these value components rely on indirect mechanisms of value accrual to determine trade-offs in final

outcomes, health system costs or both, the balance of which determines the net value of the technologies.

For the subset of studies considered most relevant for the conceptualisation, we synthesised narratively the following types of evidence:

i. Key components of value, i.e., ways in which the diagnostic technologies may lead to impacts on individuals' health and/or system cost compared to their alternatives (i.e., the comparators);
ii. Characterisation of the modelling/evidence-linkage approaches used to quantify the key indirect components of value, identifying underlying structural assumptions;
iii. Value drivers, i.e., factors expected to have a considerable impact on cost-effectiveness;
iv. Main areas of uncertainty and evidence scarcity, as well as approaches taken to deal with these issues;
v. Sources of heterogeneity, and approaches taken to handle heterogeneity;
vi. Data sources relevant to the UK decision making context.

The focus of the narrative synthesis was placed on the characterisation of value accrual mechanisms that may be relevant to the current assessment of software fusion biopsy systems, rather than exhaustive characterisation of all value components.

5.4 Methodology of the review of economic evidence provided by the companies

We reviewed the economic evidence submitted by the companies in response to requests for information (RFIs) by NICE and the EAG. We listed this economic evidence grouped into three categories:

1. Full economic evaluations that consider both costs and consequences (including cost-effectiveness, cost-utility and cost-benefit analyses);
2. Resource use and cost data;
3. Other.

Full economic evaluations were considered for inclusion in one of the two other economic reviews (Sections 5.2 or 5.3) as appropriate given their study characteristics.

Resource use and cost data were considered for the parametrisation of the *de novo* model.

5.5 Results of the review of the cost-effectiveness of MRI fusion biopsy for suspected prostate cancer

5.5.1 Search and studies identified

Record from the searches described in Section 4.1 were examined to identify potentially relevant economic records. Figure 16 in Appendix 9 shows the PRISMA flow diagram for this review which details results at each stage of the review. A total of twenty-seven studies were identified as being potentially relevant to the assessment of cost-effectiveness of software fusion biopsy vs. cognitive fusion biopsy. After screening the titles and/or abstracts, twenty-six studies were excluded. One full text publication was retrieved and assessed for inclusion, Pawha et al. (2017). This study[127] met the full set of inclusion criteria and was included in this review of software fusion biopsy for suspected prostate cancer.

We note that the economic evidence submitted by the companies in response to information requests (RFIs) by NICE and the EAG largely consisted of resource use and cost data (mostly acquisition, maintenance, and training costs) on the software fusion they commercialise. This evidence was considered for the parameterisation of the model and is discussed as appropriate in Section 6.3.7.

In addition to this, KOELIS and Kebomed also submitted economic evidence consisting of:

- A cost-analysis in a Japanese setting;
- Two business case analysis;
- A slide set describing what is referred to as a cost-benefit analysis comparing MRI-influenced biopsy using KOELIS Trinity with TRUS-guided biopsy in the US health care setting.

This evidence is not considered further in this report, as the economic analyses did not comply with the inclusion criteria of this review. For example, the cost-benefit analysis presented in the slide set did not appear to include HRQoL outcomes (only cost and diagnostic outcomes). Furthermore, the evidence provided lacked sufficient detail to be informative for the model parameterisation (e.g., the methodology, sources of evidence and assumptions were not clearly described in the business case analyses) and it was not peer-reviewed.

5.5.2 Review of Pahwa et al., 2017

The Pahwa et al. (2017)[127] study is summarised in Table 31. The quality assessment of this study is reported in Appendix 9.

Table 31 Summary of cost-effectiveness study of Pahwa et al. (2017)[127]

Study country, perspective	Population	Population characteristics	Diagnostic strategies	Analytical approach, time horizon	Outcomes
US, not stated	Biopsy-naïve men with indication for biopsy due to elevated PSA levels or clinically significant DRE findings	Mean age 65 years PCa prevalence 50% Prob CS PCa| PCa 50%	1. Systematic TRUS biopsy for all 2-4. Non-contrast mpMRI for all followed by MRI-influenced biopsy (2. cognitive fusion, 3. MRI fusion, or 4. in-bore) for those with clinically suspect lesions on mpMRI. Those without mpMRI detected suspicious lesions do not receive biopsy. 5.-7. Non-contrast mpMRI followed by MRI-influenced biopsy (5. cognitive fusion, 7. MRI fusion, or 7. in-bore) for those with clinically suspect lesions on mpMRI. Those without MRI detected lesions receive systematic TRUS biopsy.	Cohort decision tree model Life-time horizon	Costs QALYs NHB ICER

CS, clinically significant; DRE, digital rectal examination; ICER, incremental cost-effectiveness ratio; MRI, magnetic resonance imaging; NHB, net health benefit; PCa, prostate cancer; PSA, prostate-specific antigen; QALY, quality-adjusted life year; TRUS, transrectal ultrasound.

Pahwa et al. (2017)[127] evaluated the cost-effectiveness of mpMRI followed by MRI-influenced biopsy using alternative MRI-influenced methods (software fusion, cognitive fusion and in-bore MRI biopsy) compared to systematic TRUS biopsy in individuals with suspected prostate cancer in the US healthcare system. The study's perspective is not explicitly stated, but the costs included suggest a societal perspective.

The study population consisted of biopsy-naïve men with elevated PSA levels and/or clinically significant DRE findings. In the base-case analysis, the cohort had a mean age of 65 years, and a prostate cancer prevalence of 50%; this prevalence estimate was varied in subgroup analyses by age groups. Cancer prevalence by age was sourced from a study which reviewed US cancer statistics and autopsy data; it is unclear if this estimate is reflective of a biopsy-naïve population. The probability that prostate cancer is clinically significant cancer (defined as tumour volume greater than 0.5 cm^3, a Gleason score higher than 6, or with a Gleason pattern of 4 or 5 [if Gleason score ≤6] or not confined to the prostate) was assumed to be 50%, based on a previous cost-effectiveness study.

The study compared three diagnostic strategy types with the following test sequences: i) systematic biopsy for all individuals, ii) mpMRI for all individuals followed by MRI-influenced biopsy for those with clinically suspicious lesions detected on mpMRI (positive mpMRI) and no further testing for those with negative MRI findings, and iii) mpMRI for all individuals followed by MRI-influenced biopsy for those with a positive mpMRI result and TRUS systematic biopsy for those with a negative mpMRI result. Each strategy type with an MRI-influenced component (ii. and iii.) was evaluated separately for each alternative MRI-influenced method (software fusion, cognitive fusion and in-bore MRI biopsy). Individuals who did not undergo biopsy or had a negative result did not receive treatment. Those who undergo biopsy and have a positive result are classified according to cancer significance and receive treatment consisting of a mix of active surveillance, watchful waiting, androgen deprivation therapy, and radical treatments (radiotherapy, brachytherapy and radical prostatectomy). mpMRI was described as non-contrast and biopsy as TRUS; no further details on the specifications of the test were provided.

The decision model consisted of a cohort decision tree structure which characterised diagnostic pathways, treatment allocation and assigned lifetime payoffs by classification and treatment allocated. It started by classifying individuals according to their true disease status including clinical significance (no prostate cancer, clinically non-significant or significant prostate cancer). Individuals were subsequently classified according to the diagnostic accuracy of test sequences in each strategy according to diagnose results and their true underlying disease status (including disease significance).

The metrics of diagnostic accuracy for the different biopsy approaches included the sensitivity to detect a) cancer (for systematic biopsy only), and b) clinically significant cancer (only for targeted

biopsies), c) clinically insignificant cancer (only for targeted biopsies) and c) a probability of correctly identifying the tumour aggressiveness. In addition, all biopsy approaches were assumed to be 100% specific to detect prostate cancer. The diagnostic accuracy of software fusion biopsy was not reported as specific to any particular MRI-fusion technology. The evidence used to inform the sensitivity of software fusion to detect clinically insignificant cancer was pooled from various MRI-fusion systems (including Urostation, EM-tracking, MyLab Navigation [Esaote], UroNav, ei-Nav/ Artemis [Eigen]), while the sensitivity to detect clinically nonsignificant cancer was informed by evidence on Artemis™ ProFuse.

The costs considered in the model included diagnostic procedure costs (MRI), biopsies (systematic, cognitive fusion, software fusion, or in-gantry), histopathological evaluation costs, cost of workdays lost, costs of biopsy complications and lifetime treatment costs (cost-payoffs). The cost of software fusion (mean US$ 731 including physician fees) applied in the model was not technology specific.

The model does not consider the impact on HRQoL of biopsy complications.

Treatments considered in the model included radical prostatectomy, external beam radiation therapy, brachytherapy, androgen deprivation therapy (ADT), active surveillance, and watchful waiting. Treatment distributions conditional on diagnosed clinical significance were sourced from a US registry and supplemented by assumptions.

The QALY pay-offs at each terminal node are conditional on the cancer presence (and its clinical significance), treatment status (treated, untreated), and type of treatment (independent of the clinical significance of cancer). The lifetime QALY pay-offs for treated patients are mostly derived from a previous cost-effectiveness study[128] which used a state transition Markov model to compare expectant management (active surveillance or watchful waiting) with initial treatments (brachytherapy, intensity-modulated radiation, radical prostatectomy) on men with low-risk, clinically localised prostate cancer. The studies pooled to inform the treatment effectiveness in the external model are not clearly described. The Markov model captures the progression risks (from low-risk to intermediate-risk, development of metastatic prior to treatment, development from biomedical recurrence to metastatic disease), recurrence risk, short- and long-term adverse events from treatment choice on lifetime QALEs. The Pahwa et al. (2017)[127] model does not capture the probability of developing new cancer during the lifetime for men with no cancer.

The cost pay-offs are conditional on the diagnostic status (diagnosed, undiagnosed/ later diagnosed), treatments received (for diagnosed patients), and the clinical significance of cancer (for undiagnosed or later diagnosed patients). The life-time costs are also derived from Hayes et al. (2013) model.[128] As

the risk of developing new cancer is not considered, no life-time cost is assigned to men with no cancer.

Cost-effectiveness results are expressed in the net health benefit framework at a cost-effectiveness threshold of $50,000 US per additional QALY. Sensitivity analysis included probabilistic sensitivity analysis (probability distributions for each parameter not reported), one-way sensitivity analysis for all study parameters, and scenario analysis. The scenario analysis considers the cost-effectiveness of each strategy for three different levels of Gleason cut-off scores for the thresholds of clinically significant cancer: Gleason 3 + 4 or higher, Gleason 4 + 3 or higher, Gleason 8 or higher. The authors also present subgroup analysis by three age subgroups (41-50 years; 51-60 years and 61-70 years), with prevalence and life expectancy varying across subgroups.

5.5.2.1 Pahwa et al., 2017 cost-effectiveness results

The cost-effectiveness base-case results are summarised in Table 32. Strategy 4, consisting of mpMRI followed by in-bore biopsy for those who test positive on imaging and no further biopsy for those with a negative imaging result, had the highest NHB at $50,000 USD per additional QALY.

Strategies with cognitive fusion components (2 and 5) have higher NHB than the corresponding strategies with software fusion biopsy (3 and 6). than those of MRI-influenced fusion biopsy in both the base-case analysis and for scenario analysis where the definition of clinically significant disease is varied. Software fusion biopsy generally results in lower total QALYs and higher total costs compared to cognitive biopsy.

Table 32 Summary of cost-effectiveness results in Pahwa et al., 2017

	Total Costs (US$)	Total QALYs	ICER (US$ per QALY)	NHB (QALYs)* (95% CI)
Strategy 2: mpMRI, cognitive fusion biopsy, no systematic biopsy if negative	17630	9.250	-	8.997 (7.34, 10.21)
Strategy 4: mpMRI, in-bore biopsy, no systematic biopsy if negative	17870	9.308	$4147	8.950 (7.54, 10.21)
Strategy 3: mpMRI, software fusion biopsy, no systematic biopsy if negative	18608	9.198	Dominated	8.826 (7.33, 10.19)
Strategy 5: mpMRI, cognitive biopsy, systematic biopsy if negative	18802	9.269	Dominated	8.893 (7.45, 10.18)
Strategy 7: mpMRI, in-bore biopsy, systematic biopsy if negative	19042	9.326	$65111	8.946 (7.60, 10.17)
Strategy 6: mpMRI, software fusion biopsy, systematic biopsy if negative	19780	9.217	Dominated	8.822 (7.43, 10.16)
Strategy 1: Systematic biopsy	19133	9.082	Dominated	8.699 (7.08, 10.15)

*at $50 000 US per additional QALY

CI, confidence interval; ICER, Incremental Cost-Effectiveness Ratio; NHB, net health benefit; QALY, quality-adjusted life-year.

The authors claimed that the one-way sensitivity analysis results suggest that the cost-effectiveness drivers are cancer prevalence, the proportion of clinically significant cancer, and the sensitivity of MRI. However, we note that results are not presented and that the ranges within which the model parameters were varied do not seem to follow any other rationale other than assuming great parameter uncertainty and testing extreme input values. Scenario and subgroup analysis results were consistent with those of the base-case analysis.

5.5.2.2 Generalisability and relevance of the Pahwa et al., 2017 study to the decision problem in the current assessment

The Pahwa et al. (2017)[127] study has several features that limit its generalisability and relevance to the decision problem in the current assessment.

Firstly, the study's perspective does not correspond to the NICE reference case, as it seems to take a US societal perspective rather than that of NHS and personal and social services (PSS). This difference in perspective implies that the opportunity costs considered in Pahwa et al. (2017)[127] are unlikely to be comparable to those relevant to this assessment. It also means that the range of included costs in Pahwa et al. (2017)[127] are not directly generalisable to this assessment.

Another area where there is a lack of alignment between this assessment and Pahwa et al. (2017)[127] is the study population considered and how this links to the position of the tests in the diagnostic pathway. Since the study predates the routine use of MRI to screen individuals with suspected prostate cancer for biopsy, the study population is not limited to individuals with an MRI Likert or PI-RADS score equal or greater than three. The study population is also limited to those individuals without a prior biopsy. Population characteristics such as prevalence, a cost-effectiveness driver in Pahwa et al. (2017)[127] are, therefore, likely to differ between this study's population and the population defined by the scope of this assessment, thus limiting the generalisability of the study findings to this assessment.

The diagnostic pathway in the study also differs from the one currently recommended in UK clinical practice, as it does not allow for repeat biopsies.

The way in which diagnostic accuracy was modelled in Pahwa et al. (2017)[127] is another limitation, as the tests classified individuals according to prostate cancer presence and its clinical significance. Clinical recommendations for management of prostate cancer in the UK are made based of prognostic risk (characterised via a five-tier risk score), rather than clinical significance of disease alone. Therefore, the diagnostic classification in the study is insufficiently granular to allow linking classification to clinical management choices in the UK context.

Another issue in Pahwa et al. (2017)[127] is that it did not model a specific software fusion technology. The way in which the direct costs and diagnostic accuracy of software fusion were modelled, implies that these estimates are equivalent across different technologies. This assumption is not justified, but the equivalence of the direct costs of alternative technologies is debatable, even if diagnostic accuracy can be assumed equivalent given the similar functioning of these software systems. The study also does not model or discuss potential diagnostic accuracy and/ or cost modifying factors, such as the method of estimation (rigid vs. elastic), the biopsy sampling method (targeted alone vs. combined), the biopsy approach (transperineal vs. transrectal, local anaesthesia vs. general anaesthesia), etc. These factors have been identified in the scope of this assessment as features of interest and may impact on the cost-effectiveness results.

Finally, the evidence linkage between clinical management and final outcomes in the Pahwa et al. (2017)[127] model lacks flexibility to allow adaptation to other jurisdictions, since these outcomes are modelled as pay-offs estimated from an external US Markov model. It is unclear whether the distribution of treatments used to weigh the costs and QALYs pay-offs in the study is likely to match what is observed in a UK setting. However, even if the treatment distribution was reflective of UK clinical practice, the external Markov model also quantifies lifetime outcomes specific to the US setting. Therefore, it is not possible to easily implement alternative UK relevant treatment choices and reflect the impact of these on long-term cost and HRQoL outcomes.

Therefore, the EAG concludes that the Pahwa et al. (2017)[127] study cannot directly inform or be adapted to inform the decision problem in the current assessment.

5.6 Results of the additional targeted reviews to support model conceptualisation

The results of the searches are detailed in detail in Appendix 1. In total, fifteen cost-effectiveness models[126, 129-143] were considered potentially relevant to inform the *de novo* model conceptualisation for inclusion. These studies are summarised in Table 82, Appendix 9.

Of the fifteen cost-effectiveness models identified at the first stage of the review, nine were selected for a more in-depth review, as these were identified as the most appropriate to support the conceptualisation of the *de novo* model given the relevance of:

- The comparisons and position in the diagnostic pathway –studies which compared biopsies conducted with MRI-influence methods (i.e., targeted and/or combined biopsies) for prostate cancer diagnosis;[129, 130, 134, 139, 140]
- UK policy relevance.[126, 131, 133, 135, 136]

5.6.1 *Studies included in the model conceptualisation review*

Table 33 summarises the subset of identified studies included in the model conceptualisation review. A summary description of these studies is provided next; further detail can be found in Appendix 9.

Table 33 Studies included in the model conceptualisation review

Study: 1st author, year, country Type of model	Population	Biopsy approaches modelled	Classification (via biopsy diagnostic accuracy)	Choice component	Evidence linkage to longer-term outcomes	
					PCa	No PCa
Souto-Ribeiro (2022)[126], UK Diagnostic	Main population: Biopsy naïve individuals with mpMRI Likert3+ for suspected localised PCa. .Other populations: biopsy-naïve mpMRI Likert 1,2; previous negative biopsy and mpMRI Likert3+; previous negative biopsy and mpMRI Likert 1,2	LATP vs. LATRUS vs. GATP biopsy Repeat biopsy: with LATRUS for a proportion of those diagnosed as No PCa or CNS PCa (max: 1)	No PCa CNS PCa CS PCa	-- No PCa: discharge if TN; PSA monitoring if FN -- CNS PCa: either AS or radical treatment -- CS PCa*: . intermediate risk: offered radical treatment, with option of AS; %WW (if no curative intent) . high risk: % Radical treatment; %WW (if no curative intent) . Metastatic PCa: ADT±Chemo	Intermediate outcome: disease progression to metastatic disease – varies by underlying true risk category and being diagnosed as having CS or CNS PCa Survival: Metastatic disease, diagnostic status of metastatic disease, age HRQoL: Metastatic disease, age, AEs from treatment Costs: Disease spread, age, diagnostic status, treatment received, EoL	Surv: Age HRQoL: NR Costs: Monitoring
Wilson (2021)[131], UK Diagnostic	Individuals with suspected PCa presenting for mpMRI	LATP vs. LATRUS biopsy Repeat biopsy: all diagnosed no PCa at previous biopsy (max: 1)	No PCa CNS PCa CS PCa	-- No PCa: discharged back to 1ary care -- CNS PCa: AS -- CS PCa*: intermediate or high risk: AS or radical prostatectomy	Intermediate outcome: disease progression to metastatic disease – varies by underlying true risk category and treatment received Surv: Metastatic disease, age HRQoL: Metastatic disease, age Costs: Treatment received	Surv: Age HRQoL: Age Costs: NR
Cheng (2021)[129], Singapore Diagnostic	Biopsy-naïve individuals with elevated PSA level &/or abnormal DRE findings	Combined vs. systematic(12-core) vs. saturation (20-core) biopsy Repeat biopsy: all diagnosed no PCa at previous biopsy (# of repeat biopsies is strategy dependent, max 2)	No PCa CNS PCa CS PCa	-- No PCa: monitoring -- CNS PCa: AS, WW or radical treatment -- CS PCa*: intermediate or high risk: WW or radical treatment. WW only offered if no curative intent	Intermediate outcome: disease progression to metastatic disease – varies by underlying true risk category and diagnostic status Surv: Metastatic disease, age HRQoL: Metastatic disease, castration-resistant disease, age, treatment, underlying true risk category Costs: Metastatic disease, castration-resistant disease; treatment received, EoL	Surv: Age HRQoL: Age Costs: Monitoring

Hao (2021)[130], Sweden Screening + diagnostic	Men eligible (55-69 years old) for quadrennial PSA screening of PCa	Targeted biopsy vs. systematic biopsy vs. combined biopsy Repeat biopsy: not modelled as part of the diagnostic component	ISUP GG0 ISUP GG1 ISUP GG≥2	--ISUP GG0: return to screening --ISUP GG1 and GG2+: AS or radical prostatectomy and/or radiation therapy Metastatic PCa: metastatic drug treatment Treatment allocation also seems to consider disease stage at diagnosis (T1-T2, T3-T4).	Intermediate outcome: disease progression to metastatic disease – varies by underlying ISUP GG and T stage and diagnostic status Surv: Metastatic disease, other factors NR HRQoL: Metastatic disease, age, treatment and time since treatment initiation received, being diagnosed, EoL Costs: Treatment received, EoL	Surv: NR HRQoL: NR Costs: NR
NICE (2019)[133], UK Diagnostic	Individuals with raised PSA, negative MRI and/or a previous negative prostate biopsy	TPMB vs. TRUS Repeat biopsy: no consecutive biopsies allowed	No PCa CNS PCa CS PCa	-- No PCa: monitoring (tests and testing schedule differ across strategies) -- CNS or CS PCa*: mix of AS, brachytherapy, hormone therapy, radical prostatectomy, external radiotherapy with the distribution of treatments varying by underlying category of risk (low, intermediate or high risk). --Metastatic PCa: ADT±Chemo	Intermediate outcome: disease progression to metastatic disease– varies by underlying true risk category and being diagnosed as having CS or CNS PCa Surv: Metastatic disease, diagnostic status of metastatic disease, age HRQoL: Metastatic disease, age, AEs from treatment Costs: Disease spread, age, diagnostic status, treatment received, EoL	Surv: Age HRQoL: Age Costs: NR
Faria (2018)[135] & Brown (2018)[136], UK Diagnostic	Biopsy-naïve individuals with suspected localised PCa	TRUS vs. TPMB Repeat biopsy: who receives it (No PCa or CNS PCa) varied by strategy (max 1)	No PCa CNS PCa CS PCa	--No PCa: follow-up 1^{ary} care --CNS PCa: AS --CS PCa: intermediate or high-risk radical prostatectomy	Intermediate outcome: disease progression to metastatic disease – varies by underlying true risk category and treatment received Surv: Metastatic disease, age HRQoL: Metastatic disease, age Costs: Treatment received	Surv: Age HRQoL: Age Costs: NR

Barnett (2018)[134], US Screening + diagnostic	Men eligible (55-69 years) for annual PSA based screening of PCa	TRUS systematic vs. TRUS MRI fusion vs. TRUS combined biopsy Repeat biopsy: not modelled in the diagnostic component	No PCa CNS PCa CS PCa	--No PCa: routine screening --CNS PCa: if Gleason score≤6 - % AS, % radical prostatectomy; --CS PCa*: if Gleason score≥7 – radical prostatectomy; if PSA>20 ng/mL or a Gleason score ≥8 - bone scan and a CT scan for staging -- PCa (CNS & CS) & age> 80 years: WW	Intermediate outcome: disease progression to metastatic disease– varies by treatment received and indirectly by location of disease (organ confined vs. extraprostatic or with lymph node) Surv**: Metastatic disease, age, HRQoL: Metastatic disease; being diagnosed; treatment received and time since treatment initiation time post radical prostatectomy, EoL Costs: Disease spread, treatment received, EoL by age	Surv: Age HRQoL: Age Costs: Monitoring
Pahwa (2017)[139], US Diagnostic	Biopsy-naïve patients with elevated PSA level/ abnormal DRE findings. Subgroups: 41-50, 51-60, 61-70 years old	Systematic TRUS, targeted cognitive fusion, targeted MRI fusion, targeted MRI in-bore. Repeat biopsy: not modelled	No PCa CNS PCa CS PCa	-- No PCa: NR -- CNS or CS PCa: mix of AS, WW, radiation therapy, brachytherapy, prostatectomy, ADT; treatment distribution varies by diagnosed clinical significance with a higher proportion of more aggressive treatment assumed for CS PCa	Surv: Diagnostic status, age, treatment type, underlying true disease status (including clinical significance) HRQoL: Being diagnosed, age, treatment received and underlying true disease status (including clinical significance) Costs: Diagnostic status, treatment received and underlying true disease status (including clinical significance)	Surv: Age HRQoL: NR Costs: NR
Venderink (2017)[140], The Netherlands Diagnostic	Biopsy-naïve patients with elevated PSA level/ abnormal DRE findings	Systematic TRUS, targeted TRUS MRI fusion, targeted in-bore MRI biopsy Repeat biopsy: not modelled	No Pca CNS PCa CS PCa	-- No PCa: NR -- CNS or CS PCa: mix of AS, WW, radiation therapy, brachytherapy, prostatectomy, ADT; the distribution of treatments varies by diagnosed clinical significance with a higher proportion of more aggressive treatment assumed for CS PCa	Surv: Diagnostic status, treatment received, and underlying true disease (including clinical significance) HRQoL: Being diagnosed, treatment received and time since treatment initiation Costs: treatment received	Surv: Age HRQoL: NR Costs: NR

*classification for treatment allocation is not done via diagnostic accuracy alone; **not reported in full in the manuscript;

ADT, androgen depleting therapy; AS, active surveillance; Chemo, chemotherapy; CNS, clinically non-significant; CS, clinically significant; EoL, End of life; FN, false negative; FP, false positive; GATP, general anaesthesia transperineal biopsy; GG, Grade group; ISUP, International Society of Urological Pathology; LATRUS, local anaesthesia transrectal biopsy; LATP, local anaesthesia transperineal biopsy; mpMRI, multiparametric MRI; NR, not reported; PCa, prostate cancer; Surv, survival; TN, true negative; TP, true positive; TPMB, template prostate mapping biopsy; WW, watchful waiting

5.6.1.1 Scope of the study

The population in the majority of studies comprises individuals with suspected prostate cancer who enter a secondary care diagnostic pathway[126, 129, 131, 133, 135, 136, 139, 140], while other studies consider patients being screened for prostate cancer.[130, 134]

A variety of biopsy approaches were compared in the studies; these differ by route of access (transrectal vs. transperineal), type of anaesthesia used (general vs. local anaesthesia), sample collection method (targeted vs. systematic vs. mapping or saturation biopsy) and MRI-influenced methods (software fusion, cognitive fusion, and in-bore MRI). Two models are of particular interest for UK policy. Souto-Ribeiro et al., 2022,[126] reports a previous DAR by the Southampton EAG. This study established two main comparisons between biopsy approaches: i) local anaesthesia transperineal (LATP) biopsy (with any type of biopsy device) vs. local anaesthesia transrectal (LATRUS) biopsy and general anaesthesia transperineal (GATP) biopsy and ii) LATP with specific freehand devices vs. LATRUS and vs. transperineal transrectal biopsy conducted with a grid and stepping device conducted under local or general anaesthetic. The NICE CG131 model[133] evaluated alternative follow-up strategies of individuals with suspected prostate cancer and placed little emphasis on alternative biopsy approaches.

Some studies (but not all) modelled the possibility of repeat biopsies.[126, 129, 131, 135, 136] These studies varied in how they specified: who would receive a repeat biopsy, what proportion of those eligible would receive one (or more) repeat biopsies, the type of biopsy received, and the number of subsequent biopsies allowed (if more than one).

5.6.1.2 Classification

In most studies, the diagnostic accuracy of the biopsy procedure classifies individuals as not having prostate cancer or having non-clinically significant or clinically significant prostate cancer.[126, 129, 131, 133-136, 139, 140] The exception was the study by Hao et al. (2021), in which classification is done by ISUP grade.[130] Both types of classification are usually defined by histopathological features of the biopsied lesions (graded according to Gleason scores).

The specificity of biopsy to detect prostate cancer is assumed perfect across most models, so individuals without prostate cancer cannot be misclassified as having the disease. However, some studies considered the possibility of individuals with clinically non-significant prostate cancer misclassified as clinically significant;[134, 139, 140]

5.6.1.3 Choice of clinical management

Decisions on patient management at diagnosis could be determined by the biopsy diagnostic outcomes alone[135, 136, 139, 140] or with other factors also influencing treatment allocation.[126, 129-131, 133, 134]

In three models[135, 136, 139, 140] patient management was attributed according to individuals' classification in terms of disease presence and clinical significance of disease. This classification was established based on the diagnostic accuracy of the biopsy approaches. Some models, tracked the individuals underlying cancer prognostic risk and used this information jointly with the diagnostic outcomes to allocate treatment. For example, the Southampton DAR model[126] allocated treatments based on disease presence, clinical significance of disease and underlying cancer risk distribution.

For patients diagnosed with prostate cancer, the primary treatment allocation was conditional on:

i. Diagnosed clinical significance of disease, true cancer risk category and disease spread;[126, 133]
ii. diagnosed disease clinical significance;[135, 136, 139, 140]
iii. Gleason score, PSA level and age;[134]
iv. type of biopsy (targeted or systematic), cancer risk category and age.[129]

A range of evidence sources were used to inform the distribution of treatments for diagnosed prostate cancer. Amongst these the following are relevant in the UK context:

- The Southampton DAR model[126] based treatment distribution by risk category on UK clinical guidance and observed treatment allocation from national audit data.[144]
- The NICE NG131 model[133] used observed primary treatment distributions by risk category from UK registry data.[122]
- The PROMIS trial[135, 136] assumed that treatment choice was guided by diagnosed disease clinical significance alone.

Individuals diagnosed as not having prostate cancer were discharged to follow-up,[131, 133, 135, 136] or returned to the screening schedule.[130, 134] One study,[126] conditioned the individuals' subsequent management after a no prostate cancer diagnosis on whether they had been misclassified (TN results led to discharge and FN results [patients with prostate cancer of any risk category] to routine PSA monitoring). This assumption was not justified and it is not clear how in clinical practice the two groups of individuals (TN and FN) would be distinguished so that distinct treatment decisions could be made for each group.

5.6.1.4 Outcomes

The evidence linkage approaches applied in the identified studies to connect patient classification and subsequent treatment choices with longer-term outcomes differed in whether prostate cancer progression was explicitly modelled as an intermediate outcome or not.

Only two studies did not model disease progression.[139, 140] Pahwa et al. (2017)[139] conditioned lifetime QALYs and cost payoffs on diagnostic status (i.e., whether cancer had been diagnosed or remained

undiagnosed), underlying true disease status (no prostate cancer, clinically non-significant or clinically significant prostate cancer) and type of treatment received. Venderink et al. (2017)[140] used a long-term Markov model that only allowed for transitions from alive to death states, with survival conditional on type of treatment received and the underlying true disease clinical significance, with the diagnostic status (diagnosed vs. undiagnosed cancer) determining whether individuals received treatment.[140]

All other models considered disease progression from localised to metastatic disease although health states and possible state transitions varied across models.[126, 129, 131, 133-136] Some studies modelled progression from localised to metastatic disease, and conditioned disease progression on underlying risk category and being correctly diagnosed/treatment received.[129, 131, 135, 136] Other studies modelled sequential disease progression across disease risk categories (from low to intermediate-risk and from the latter to high-risk disease) for localised disease followed by progression from the high-risk localised to metastatic disease. In these models, the probabilities of transitioning to later disease stages were conditioned on the underlying true disease status (including risk category) and being diagnosed as having clinically significant or non-significant disease.[126, 133] The screening studies modelled progression differently in the preclinical stage and in the clinical states.[130, 134].

All the disease progression models shared the assumption that prostate cancer mortality only applied to patients with metastatic disease. Treatment for patients identified as having cancer reduced disease progression to metastatic cancer compared to untreated patients, and thus reduced the probability of dying from prostate cancer for these patients. The transition probabilities for treated and untreated patients in the Markov disease progression were estimated by calibration or partially observable Markov model decision processes (as progression is an unobservable process). The data sources and calibration methods used to estimate these transition probabilities differed across models, are reviewed below for the two most relevant UK models. Details on the remaining models are in Appendix ## (#crossref).

The PROMIS model[135, 136] calibrated the probability of progressing from localised to metastatic disease by risk category and treatment received, combining risk stratified survival data and proportion of patients with metastases from the Prostate Cancer Intervention versus Observation Trial (PIVOT),[119] with the mortality in the metastatic subgroup of the Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy (STAMPEDE) trial.[123] The PIVOT observation arm was used to inform the transition probabilities for individuals with prostate cancer who did not receive active treatment (due to correct classification on misclassification depending on the risk category). The PIVOT radical prostatectomy arm was used to inform the transition probabilities for those treated with active treatment (true positives with intermediate and high-risk

cancer). The 'treatment' effects of being diagnosed on disease progression were thus informed by randomised comparative efficacy evidence.

The model used in the previous DAR [126] and in the NICE NG131 model [133] disaggregated disease progression by cancer risk categories and used calibration to estimate transition probabilities. The calibration method estimated transition probabilities first for the transition from high-risk to metastatic disease, then from intermediate to high-risk disease, and finally from low-risk to intermediate-risk disease can be derived. The calibration was done separately for the undetected and detected cancers using different data sources. Transition probabilities for the undetected cancers used cumulative metastases risk rates by cancer risk category from the watchful waiting arm in the Scandinavian Prostate Cancer Group Study Number 4 (SPCG4) trial[145] jointly with and Swedish life-table data (from 1999 to reflect background mortality in the trial). For the diagnosed cancers, the data sources for calibration included: cancer-specific survival by risk category sourced from a UK registry study,[122] all-cause survival for people with metastatic prostate cancer from the STAMPEDE trial,[59] and UK life-table (from 2010-2012 to reflect background trial mortality in STAMPEDE). Thus, this calibration approach relies on an indirect naïve comparison to derive the 'treatment' effects of being diagnosed on disease progression, which may introduce bias on the probabilities of disease progression used in the model.

In general disease progression models, survival outcomes for individuals with prostate cancer were conditional on having metastatic disease and age. Two models[126, 133] further conditioned mortality on whether metastatic disease was diagnosed (and therefore, received treatment for metastatic cancer) or not. Metastatic mortality data sources of relevance to the UK context include different publications of the STAMPEDE study, a UK based trial which compared the survival outcomes of men with newly diagnosed metastatic, high risk or node positive cancer treated with alternative cancer treatments. The PROMIS and related models estimated the probability of metastatic death using early (median follow-up of 20 months) survival data of men with newly diagnosed metastatic prostate cancer from the control arm (who received standard of care consisting of androgen depleting therapy) of the STAMPEDE trial. The NICE NG131 and related models, used a later survival data cut (median follow-up 43 months) from the docetaxel and control arms of the STAMPEDE trial that includes individuals with metastatic and non-metastatic disease.[59]

HRQoL outcomes of patients with prostate cancer were most frequently conditioned on having metastatic disease[126, 129-131, 133-136], age[126, 129-131, 133-136], and treatment received and time since treatment initiation[130, 134, 140] although other factors having been considered in select models (Appendix 9). The UK relevant utility sources for patients with prostate cancer in the long-term outcome models include Torvinen et al., 2013[146] – for the disutility of metastatic disease, Ara and Brazier, 2010[147] – for the

disutility of ageing and Mowatt et al., 2013[143] – for the disutility of treatment related adverse events (combined with rates of adverse events from Donovan et al., 2016[148]).

Most models considered the cost of treatment for patients with diagnosed localised or locally advanced prostate cancer (radical treatment or active surveillance)[126, 129-131, 133-136, 139, 140] and management of treatment adverse events.[126, 131, 133, 135, 136] Patients with undiagnosed prostate cancer would incur the costs of routine follow-up[126, 129, 131, 133, 135, 136, 139] or of delayed radical treatment.[139] The studies also considered the costs of metastatic disease treatment with or without staging and follow-up tests.[126, 129, 131, 133-136] Two models assumed diagnosed metastatic disease would be treated differently if diagnosed (docetaxel would be added to androgen depleting therapy) compared to undiagnosed metastatic disease and that treatment with docetaxel would vary with age.[126, 133] Some models included an end-of-life cost for patients who died from prostate cancer,[126, 129, 130, 133, 134] with one study conditioning the end-of-life costs on age at death.[134]

The costs of individuals who did not have prostate cancer were not clearly reported for most models, but, where reported, consisted of the costs of routine follow-up.[126, 129, 133, 134]

In UK relevant models, treatment and follow-up resource use was informed mainly by UK (clinical and technology appraisal) guidance, as well as other published data (for example, a randomised control trial informed adverse event rates of treatment[148]) and supplemented with assumptions. End-of-life costs were uprated to the relevant price year based on Round et al. (2015)[149] Unit costs were sourced mainly from national published sources.

5.6.2 *Critical review*

5.6.2.1 Value components

5.6.3 Value components

The value components of the biopsy tests in the studies included in the conceptualisation review are summarised in Table 34, which distinguishes between value components that require evidence linkage and those that are direct impacts of the tests.

Table 34 Summary of biopsy value components identified in the studies

Value components requiring evidence linkage	Studies (1st author (year))	Mechanism
Improved outcomes due to increased/earlier detection of **cancer**, i.e., fewer PCa classified as no PCa	Souto-Ribeiro (2022)[126]; Wilson, (2021)[131]; Cheng (2021)[129]; Hao (2021)[130]; NICE (2019)[133]; Faria (2018)[135]/Brown (2018)[136]; Barnett (2018)[134]; Pahwa (2017)[127]; Venderink (2017)[140]	via diagnostic accuracy identifying true cancer status and treatment outcomes
Reduction of undertreatment: improved outcomes due to increased/earlier detection of **CS PCa**, i.e., fewer CS PCa treated as CNS PCa	Souto-Ribeiro (2022)[126]; Wilson (2021)[131]; Cheng (2021)[129]; Hao (2021)[130]; NICE (2019)[133]; Faria (2018)[135]/Brown (2018)[136]; Barnett (2018)[134]; Pahwa (2017)[127]; Venderink (2017)[140]	via diagnostic accuracy and assumptions on treatment distribution and impact of treatment on outcomes, which is conditioned on true clinical significance of PCa, true cancer risk category or cancer grade
Reduction in overtreatment: improved outcomes due to improved detection of **CNS PCa**, i.e., fewer CNS PCa treated as CS PCa	Barnett (2018)[134]; Pahwa (2017)[127]; Venderink (2017)[140]	via diagnostic accuracy and assumptions on treatment distribution and impact of treatment on outcomes, which is conditioned on true clinical significance of PCa
Change the number of repeat biopsies with impacts on biopsy costs and adverse events	Souto-Ribeiro (2022)[126]; Wilson (2021)[131]; Cheng (2021)[129]; Faria (2018)[135]/Brown (2018)[136]	via diagnostic accuracy and decision rule on which individuals are eligible for a repeat biopsy
Value components with direct impacts		
Biopsy procedure costs	Souto-Ribeiro (2022)[126]; Wilson (2021)[131]; Cheng (2021)[129]; Hao (2021)[130]; NICE (2019)[133]; Faria (2018)[135]/Brown (2018)[136]; Barnett, (2018)[134]; Pahwa (2017)[127]; Venderink (2017)[140]	-
Harms and/or costs of biopsy adverse events	Souto-Ribeiro (2022)[126]; Wilson (2021)[131]; Cheng (2021)[129]; NICE (2019)[133]; Faria (2018)[135]/Brown (2018)[136]; Barnett (2018)[134]; Venderink (2017)[140]	-

Direct value components of biopsy included the costs of the procedure, and its adverse events (with associated complication costs and negative health impacts).

The indirect value components identified here are linked to diagnostic accuracy.

All studies in the conceptualisation review modelled two common value components requiring evidence linkage to be quantified; these are: an improvement of outcomes resulting from an increased and/or earlier detection of prostate cancer and of clinically significant prostate cancer. To capture the value of increased/earlier detection of clinically significant prostate cancer, the majority of models determined a single clinical management strategy for each biopsy classification option. Classification (under an assumed clinical management strategy), together with true disease status (either true cancer risk category, e.g., NICE CG131 model[133]), or cancer grade, e.g., Hao et al., 2021[130]) was then linked to to outcomes. Clinical management strategies either consisted of a single treatment option[135, 136], or a particular mix of treatments [133]).

Only three studies explicitly modelled the impact on outcomes resulting from improved detection of clinically non-significant prostate cancer.[134, 139, 140] Although the evidence linkage requirements for modelling this value component are similar to those described above for the increased and/or earlier detection of clinically significant prostate cancer, these are the only models in which the parameterisation of biopsy diagnostic accuracy allowed for clinically non-significant prostate cancer to be misclassified as clinically significant. Individuals who have been misclassified thus incur the costs and harms of unnecessary radical treatment but have limited ability to benefit in the long-term from treatment, compared to those who have clinically significant disease.

Another value component relates to the costs and/or harms incurred for individuals who undergo a repeat biopsy conditional on the result of the index (or subsequent to index) biopsy. Although these costs and harms are a direct impact of the biopsy, this is classified as an indirect value component because the decision to repeat the biopsy is conditional on the classification of the index biopsy in the testing strategy and, therefore, requires evidence via linkage. Differences in diagnostic accuracy between biopsy approaches partially determine the proportion of individuals classified as eligible for a repeat biopsy, i.e., the proportion of those who will incur the costs and harms of an additional biopsy. In addition to the linkage via classification, modelling this value component also requires a decision rule to define who is eligible for a repeat biopsy (e.g., all or a proportion of the individuals classified as not having clinically significant cancer at the previous biopsy in the test sequence). One study further assumed (in scenario analysis only) that with one type of biopsy a smaller proportion of individuals initially classified by the previous biopsy in the test sequence as eligible for a repeat biopsy would receive repeat biopsies compared to the alternative biopsy approach.[126]

The biopsy value components with direct impact on outcomes modelled in the studies included the costs of the biopsy procedure, and the costs of managing adverse events of biopsy, as well as the detrimental health impacts of adverse events.

5.6.4 Evidence linkage

The evidence linkage used to model the indirect value components relied in most studies on a common model structure whereby a decision tree approach to track individuals' diagnostic outcomes (and, in some models, biopsy adverse events) was linked to a Markov model to capture long-term outcomes.

In most models, diagnostic classification categorised individuals (correctly or not) as having i) no prostate cancer, ii) clinically non-significant or iii) clinically significant prostate cancer. The definition of clinical significance differed across models but was generally defined in terms of a Gleason score threshold or a three-tier cancer risk categorisation (defined in terms of Gleason score, PSA levels and cancer stage). This stratification reflects differences in diagnostic accuracy and prognostic for individuals in the different risk categories. In general, the low-risk disease category was assumed to correspond to true non-clinically significant prostate cancer, while the intermediate and high-risk cancer categories corresponded to clinically significant disease.

Treatment allocation for each diagnostic classification group was usually determined. This could be a single treatment option for each group (such as in PROMIS[135, 136] where all of those identified with clinically significant cancer received radical treatment). Or it could be a pre-defined mix of treatments, where the distribution of treatments differs by group (e.g. with a higher proportion of radical treatments for those at higher cancer risk)[133]). In either case, the linkage does not aim to disentangle the outcome for the diagnosed/treated by treatment received).

In most studies, the impact of being correctly or incorrectly classified by the biopsy was modelled as an effect on disease progression to metastatic cancer, and prostate cancer death only affected individuals who were in metastatic disease health states.

6 INDEPENDENT ECONOMIC ASSESSMENT: YORK MODEL

6.1 Diagnostic strategies

The model evaluated two strategies for two alternative comparisons: i) targeted software fusion biopsy vs. targeted cognitive biopsy and ii) combined (targeted and systematic) software fusion biopsy vs. combined cognitive biopsy. The four strategies could not be incrementally compared due to the mechanism of evidence generation for the diagnostic accuracy, which relied on separate evidence networks.

The test sequence and clinical management for each strategy:

1. All patients receive the index biopsy;
   1.1. if biopsy result suggests no prostate cancer or ISUP grade 1, a proportion of patients undergoes repeat biopsy. Patients who do not undergo repeat biopsy are managed in accordance to their diagnosed ISUP grade/ CPG or discharged to routine monitoring.
   1.2. If biopsy result suggests ISUP grade 2 or greater, the individual receives treatment according to CPG.
2. For the patients who receive repeat biopsy;
   2.1. Individuals are clinically managed according to the highest ISUP grade/CPG score between the two biopsy results or discharged to routine monitoring if the biopsy suggests no prostate cancer.

6.2 Model development

6.2.1 Conceptualisation

The value components identified in the review supporting conceptualisation (Section 5.6) were:

- direct value components of biopsy, including the costs of the procedure and its adverse events (with associated complication costs and negative health impacts), and,
- indirect value componets, including the increased or earlier detection of any prostate cancer, of clinically significant prostate cancer, or of non-clinically significant cancer, and the reduction of repeat biopsies.

From the review supporting the conceptualisation (Section 5.6), we have identified several key aspects to consider in the conceptualisation of the *de novo* model, which we describe below and pertain to the diagnostic accuracy, the concept of under and overdiagnosis, the modelling of disease progression and issues with outcome evidence sources.

The histopathological biopsy results are expressed in terms of Gleason score (see Section 2.1) and sometimes including lesion core length or cores positivity. In order to estimate the diagnostic accuracy measures applied in the models, the results of the biopsy are typically collapsed into one no prostate cancer and two prostate cancer categories (clinically non-significant and clinically significant). The collapse of diagnostic information into these categories implies an information loss, as the granularity of biopsy results is not preserved in the classification according to biopsy accuracy (Gleason score ranges from 2 to 10). It also implies a judgement on the definition of clinically significant disease at a specific Gleason threshold, with some models using a Gleason threshold of 3+3 and others 3+4.

Furthermore, making clinical management choices between active surveillance and a range of radical surgical treatments and/or radiotherapy requires information provided by the biopsy diagnostic accuracy, but also information with prognostic value like PSA levels and disease stage at diagnosis. In clinical practice, patient preference is also another factor influencing the choice of management strategy. Due to this, several models made assumptions on how to map from the two prostate cancer classification into three-tier risk cancer prognostic risk classifications. Current UK clinical guidance[10] for the management of newly diagnosed localised or locally advanced prostate cancer recommends an even more granular prognostic risk classification, the CPG system, which uses the same type of information as the previous risk classification but classifies patients into five categories. The most recent update of the NICE CG131 defines four alternative clinical management strategies for individuals diagnosed in the different groups (same treatment strategy for CPG 4 and 5), whereas previous guidance defined three management strategies (one for each risk category).

The concepts of under/overtreatment are not clearly defined in the literature. In general terms, overtreatment seems to arise when patients with prostate cancer of favourable prognostic receive radical treatment (e.g., radical prostatectomy or radiotherapy) instead of active surveillance. In contrast, undertreatment would arise when patients with worse disease prognosis receive active surveillance, rather than radical treatment. So undertreatment/overtreatment can occur if the clinical management approach taken is not commensurate with the true disease prognostic risk, which may be due to:

i. disease not been correctly classified in terms of its underlying prognostic risk; and/or
ii. the prognostic risk categorisation not being accurately predictive; and/or
iii. treatment decision rules not being followed due to clinical variation and/or patient preference.

The move from the three-tier to the five groups classification, aims to improve the identification of patients who have slow progressing disease and should be managed with active surveillance. For these patients the harms (and costs) of radical treatment are likely to offset its long-term benefits.

The misclassification of individuals in the lower risk categories/groups as having a higher prognostic risk (overdiagnosis) may result in net health losses if it leads to unnecessarily radical treatment (overtreatment). Therefore, reducing overtreatment is an important value component of biopsy. The few previous studies which modelled this value component did so by capturing misclassification of clinically non-significant as significant cancer and linking this to the outcomes of more radically treated patients. This is an imperfect link, as it lacks the flexibility to identify individuals with clinically significant who are at the lower end of the prognostic risk spectrum (i.e., CPG 2 or favourable intermediate risk), and, thus, quantify the net benefit of providing active surveillance to this group.

Most studies modelled the reduction of underdiagnosis, i.e., the value of increased or earlier detection of prostate cancer in individuals whose disease will progress at a faster rate if not managed with radical treatment. This value component was modelled by capturing misclassification of clinically significant as non-significant cancer (or no cancer) and linking this to the outcomes of patients undiagnosed for clinically significant cancer. Since this classification does not allow the identification of individuals with favourable intermediate risk, it may overestimate the net benefit of treating with more radical treatment individuals with true clinically significant prostate cancer.

While most studies modelled longer-term outcomes as a function of prostate cancer disease progression, we identified two alternative structural choices to model the unobservable disease progression: i) directly between localised (or locally advanced disease) to metastatic disease, and ii) sequentially progression across three health states defined by category of true underlying prognostic risk. These two approaches also differ in terms of evidence requirements for parameterisation, with the second approach requiring more data and/or more structural assumptions to be imposed in the model. We also identified alternative methods to estimate unobservable transitions probabilities, namely calibration and partially observed Markov process models.

We have also identified issues with outcomes evidence. Some models used naïve/unadjusted comparisons, i.e. used different data sources to describe outcomes for different groups. This may result on bias. Additionally, all models used data sources to describe ouctomes according to true disease that use an imperfect reference standard (typically PSA results).

These key aspects grounded the *de novo* model conceptualisation, an overview of which is provided below.

Risk stratification: In terms of risk stratification, and given that the current UK clinical guidance[10] recommends a five category prognostic risk classification, the CPG system, there is the need to consider this more granular classification system in the modelling. Despite this being a 5-tier classification system, only four alternative clinical management strategies are recommended in the NICE guideline (same treatment strategy for CPG 4 and 5), therefore CPG4-5 can be reasonably collapsed in analysis. However, broader evidence does not typically use the CPG system, for example, we found no diagnostic studies reporting results using CPG, and therefore ISUP grade was used in the diagnostic component to reflect CPG tiers.

Determining diagnostic accuracy: The review work (Section 4.1) focused on identifying and synthesising studies (RCTs and within-patient comparisons) comparing cognitive fusion and software fusion targeted prostate biopsy methods. The multinomial model used in the synthesis of this evidence (Section 4.2.2) compares the alternative biopsy methods in how they classify individuals across the

following categories: 1 (no prostate cancer), 2 (ISUP Grade 1), 3 (ISUP Grade 2), 4 (ISUP Grade 3) and 5 (ISUP Grade 4 or 5 pooled together). This allows a more complete consideration of evidence across ISUP Grades, extending from previous approaches that focus on either cancer detection rates (typically defined as no cancer vs. ISUP Grade ≥ 1) or detection rates of clinically significant cancer (typically defined as no cancer or ISUP Grade 1 vs. ISUP Grade 2 or above).[126, 127, 129-131, 133-136, 140]

The synthesis model considers the distribution of individuals by ISUP Grades and relates this distribution across technologies using a set of odds ratios, the quantities pooled across studies. Note that such a model does not identify concordance between methods in biopsy test results (further explanation in Appendix 10). The application of the synthesised odds ratios to an externally derived distribution of probabilities of test results for one of the tests (say software fusion) retrieves the expected distribution of probabilities for the other test (cognitive fusion). This calculation of absolute probabilities is described in Appendix 10.

The evidence synthesis model does not consider the accuracy of either method in relation to a reference standard (by virtue of the evidence available for inclusion) i.e., it does not consider the extent of misclassification with either any of the modelled methods. This has important implications for economic modelling as, in the absence of a robust and representative outcomes RCT, evidence linkage is required, facilitated by knowing the extent of misclassification of the different tests in relation to true disease status, to allow determine its consequences to health and economic outcomes.

To consider accuracy evidence, a structural approach is required that extends the synthesis model to integrate such evidence. The approach developed here is described in Section 6.3.1.

Diagnostic pathway and repeat biopsy: The need and the accuracy of repeat biopsies is a potential value component for software fusion methods, in relation to cognitive fusion. This may arise indirectly from improved diagnostic accuracy of the method used for the first biopsy, i.e. a more accurate identification from a first biopsy can lead to a decreased pool of individuals eligible for re-biopsy. We did not identify comparative evidence suggesting differences in the rates of repeat biopsy between cognitive and software fusion. However, the clinical advisers to the EAG suggested that a potential value component for software fusion, is that by consulting the stored cartograms produced by MRI systems, the MDT could better target re-biopsy. There is, however, a lack of evidence to parameterise impact beyond what can be captured via diagnostic accuracy. We will explore the potential value of such a case in scenario analyses.

Treatment of prostate cancer: There is UK relevant evidence on the distribution of treatments for patients identified at different CPG groups. Our model will therefore be reflective of the different mixes of treatments used at different CPG levels (see Section 6.3.4)

Modelling of long-term outcomes: To reflect the value of increased/earlier detection, the long-term outcomes component of the model will need to condition on true disease status and the diagnosed disease category (given the prostate cancer management strategy determined by the diagnosed disease category). None of the existing long-term models have been developed using the five category prognostic risk classification based on CPG system, recommended in the current UK clinical guidelines[10]. Therefore, a *de novo* inference model will be developed for this assessment. For its structure, and given that this assessment focusses on the diagnostic pathway, considering prostate cancer disease progression over time and incidence is not as relevant as for the NG131 model, which aimed to model monitoring startegies. Therefore, the increased complexity of the structure used in the NG131 model[133] (and in the Southampton DAR[126]) may not be justified for the purpose of modelling biopsy within the diagnostic pathway. Additionally, evidence to support such a complex structure is sparse (if existing at all), and therefore its parameterisation would rely on a number of assumptions that cannot be verified. However, the added complexity of such a structure would allow for the time profile of treatment costs on those that leave the diagnostic pathway under a monitoring strategy to be better captured.

In terms of evidence to quantify the impact of alternative treatments on outcomes, comparative effectiveness evidence will be preferred to avoid bias. The most contemporary evidence available will be used to inform the inference submodel.

Further details on the inference model and on how this will be incorporated in the cost-effectiveness decision model are provided in 6.3.3.

6.3 Model structure and parameterisation

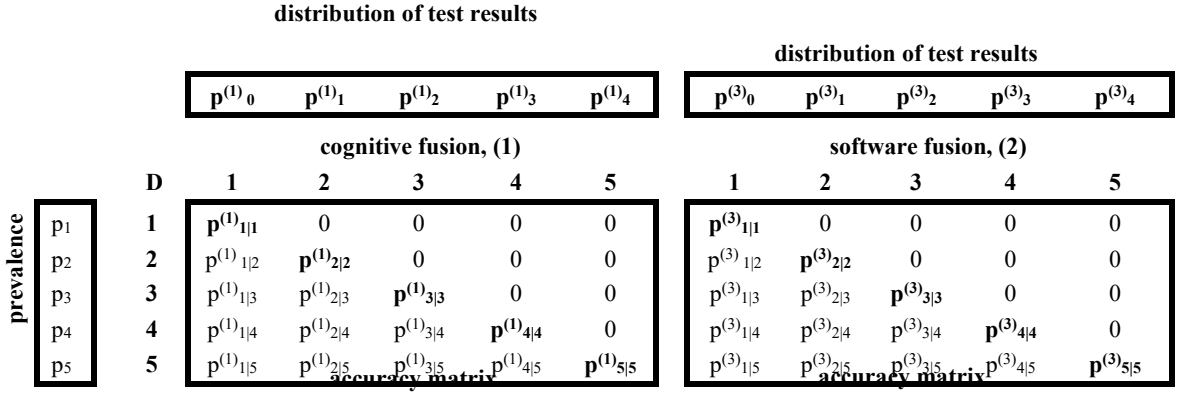
6.3.1 Modelling of first biopsy results

6.3.1.1 Determining diagnostic accuracy

As identified above (section 2.1), the fact that the evidence synthesis conducted as part of this assessment does not consider the accuracy of the different biopsy methods in relation to a reference standard has important implications for economic modelling. In the absence of a robust and representative outcomes RCT, economic modelling relies on evidence linkage facilitated by knowing the extent of misclassification of the different tests in relation to true disease status and determining its consequences to health and economic outcomes.

The extent of misclassification can, however, be made explicit by the accuracy matrix, the elements of which reflect the probabilities of obtaining a particular test result with one method conditional on a particular level of (true) disease status. Together with prevalence estimates, this matrix determines the distribution of test results, shown at the top of Figure 7.

Figure 7: Illustration of the relationship between prevalence, and the accuracy and distribution of test results across five categories, for two hypothetical tests



Note that, due to the nature of biopsy and histological examination of the biopsy specimen, it is reasonable to assume that false positive results are not possible, i.e., if cancer is histologically identified, then it is present. This implies that biopsy methods cannot identify a higher category than true disease status, and therefore zero probability is attributed to such cases in the above accuracy matrices.

Where multiple methods are of interest, the problem becomes more complex for two reasons. First, the prevalence (i.e., the true distribution across categories) is independent of test results and therefore a common prevalence estimate needs to ground all distributions of test results, and be consistent with these. Second, explicit accounts of accuracy need to respect both the prevalence estimates and the marginal distribution of test results derived from the synthesis. Therefore, a structural approach is required for determining accuracy from the marginal distributions obtained through application of the synthesis model.

Summary of approaches used in previous cost-effectiveness models

From the conceptualisation reviews (see Section 5.6), two cost-effectiveness reviews have focussed on a similar context where no accuracy evidence was synthesised.

A previous DAR[126], from now on referred as the Southampton DAR, conducted a meta-analysis on cancer detection rates (using relative risks) including studies comparing the biopsy methods of interest to the decision problem (e.g., LATP versus LATRUS), and did not include evidence comparing either method to a reference standard. In this work, the authors sourced the baseline distribution for LATRUS and its accuracy matrix, from an external diagnostic accuracy study (the PROMIS study[135,Brown, 2018 #1506]). The authors then applied the synthesised relative risks of cancer detection for LATP biopsy (derived for marginal distributions) directly to both (1) the conditional probability of LATRUS identifying clinically significant cancer conditional on true disease status, and to the (2) conditional probability of identifying clinically non-significant cancer (assumption imposed in the

base case). The conditional probability of no cancer given true disease status was then adjusted to be 1 minus the remaining. The way the relative risks were applied in the model is not consistent with the way in which they were derived, in that the relative risk derived from the synthesis model refers to the relative increase in detection rate with one method in relation to another; the relative risks were therefore derived on marginal probabilities and not on conditional probabilities. Their application to conditional probabilities in such a way implies that the increase in accuracy of detecting cancer with a particular test is independent of whether the cancer was clinically significant or non-clinically significant, and that the increase in accuracy of detecting non-clinically significant cancer given the patient has non-clinically significant cancer is equal to the increase in accuracy of detecting clinically significant cancer, given the patient has clinically significant cancer.

An alternative study, Wilson et al. (2021)[131], also investigating LATP in relation to LATRUS, assumed no difference in the expected accuracy of the biopsy methods in the comparison of interest. Therefore, the authors sourced prevalence and accuracy estimates for LATRUS from the PROMIS study and used it to represent the expected results for both biopsy methods. In reflecting uncertainty, the authors sample from the accuracy matrix directly, taking two independent samples to represent the two different biopsy methods, and therefore generate differences in the accuracy matrix between the methods, due to randomness only.

None of the existing approaches has direct applicability in the current assessment, where a disaggregation by Gleason Grade is required.

6.3.1.2 Summary of methods

The approach used in the current assessment was designed to:

- be grounded on the results of the evidence synthesis model,
- return a true distribution across ISUP Grade categories (prevalence) that is internally valid, i.e., that is not lower than the estimated ISUP Grade detection rates of the different biopsy methods,
- be grounded on available evidence on the likely accuracy of targeted MRI-fusion conditional on ISUP Grade, and
- define accuracy matrices for the remaining biopsy methods of interest that are consistent with both prevalence and the distributions of biopsy results from the evidence synthesis.

To achieve this, an extension to the synthesis model was developed, drawing on the broader evidence in Section 4.7.1. To allow for an internally consistent approach, we grounded our methodology on the distribution of test results obtained with MRI-influenced methods and their accuracy. Given that disease prevalence is fully determined by these two results, the prevalence evidence identified in

Section 4.7.1 will not be explicitly incorporated in our analyses but will instead be used qualitatively to put our results into context.

The methodology is summarised below. A more comprehensive description of the methods used is presented in Appendix 10.

1) Distribution of test results

The distributions of test results across the disease categories for the relevant biopsy methods within each disconnected component of the network in Model 1a were computed by applying network-specific baseline distributions to the results of the NMA. Building from the analyses in the evidence synthesis section, the baseline distributions were sampled from a Multinomial likelihood with an uninformative Dirichlet prior distribution for its hyperparameters, to allow for uncertainty in describing the data from the empirical studies.

2) Accuracy matrix for software fusion

Evidence on the accuracy matrix for software fusion, sourced from the literature, was used to characterise the elements of the accuracy matrix probabilistically in the model. A Multinomial likelihood was used to describe the distribution of test results conditional on each particular level of true disease status (each line in the matrix in Figure 7) with Dirichlet uninformative prior distributions.

3) Prevalence

The derivation of prevalence followed two steps, the first consisted of the analytical derivation of an initial prevalence estimate from the marginal distribution and accuracy matrix for software fusion. The second step entailed applying a constraint to ensure that the prevalence is always higher than the detection rates (by ISUP grade) observed across all tests.

4) Accuracy matrix for remaining biopsy methods

The diagonals of the accuracy matrices for the remaining biopsy methods were determined by the prevalence and the test-specific distribution of results. To define the remaining non-zero and free elements of the matrix, uninformative Beta distributions were used, constrained so that their multiplication by the prevalence retrieves the test results estimated within the evidence synthesis.

Implementation

The extension to the synthesis model developed to determine accuracy was implemented alongside the synthesis model, in a Bayesian framework estimated through Markov chain Monte Carlo methods

using WinBUGS 1.4.3.[150] Due to the sparseness of evidence in other networks, this was applied to Model 1a (Section 4.4.1.1) which includes software fusion, cognitive fusion and systematic biopsy in a first connected network, and the combination of software and cognitive fusion with systematic biopsy in a second connected network. As in the evidence synthesis, model convergence was assessed where possible by running two independent chains with different starting values looking at history plot and through inspection of Gelman-Rubin diagnostic plots. Model fit was assessed by comparing the mean total residual deviance to the number of independent data points contributing to the analysis.[71]

Sensitivity analysis

Given that the approach proposed here is heavily data driven, sensitivity analyses focussed on varying the data sources for the baseline distributions and accuracy matrix.

6.3.1.3 Results

The extension to the synthesis model reflects the data sources described in Section 4.4.1.1 for the baseline distribution of test results for software fusion, the reference method. The extension model also required data to characterise the accuracy matrix for the reference biopsy method, two sources for these data were available (see Section 4.7.1.3) and were used here. According to the data sources used, the following analyses were conducted:

- Main analysis, for the subgroup of biopsy-naïve individuals: baseline distribution of test results for software fusion sourced from biopsy-naïve data from Filson (2016)[96], relative accuracy data from the multinomial evidence synthesis model (Section 4.2.2) which was incorporated in this extension, and accuracy data from Mortezavi (2018)[107]. Mortezavi (2018)[107] was chosen for the main analysis over Zhou (2018)[108] as it more closely reflects the lower accuracy observed in UK-specific evidence sources.
- Subgroup analysis for previous negative biopsy individuals: all sources were equal to those used in the main analysis except the baseline distribution of test results for software fusion which was sourced from previous negative biopsy data from Filson (2016)[96].
- Sensitivity analysis to data source on baseline distribution: all sources were equal to those used in the main analysis except the baseline probabilities, which were based on biopsy-naïve data from PAIREDCAP (2019)[88] for network 1.
- Sensitivity analysis to data source on accuracy matrix: all sources were equal to those used in the main analysis except accuracy data which was sourced from Zhou (2018)[108].

Note that given the two networks are disconnected, results are reported separately for comparisons of cognitive fusion and software fusion – network 1, and for comparisons of combined

cognitive/software fusion with systematic biopsy – network 2. Note that while network 1 includes systematic biopsy, results for this biopsy method are not reported here.

Main analyses (biopsy naïve)

Table 35 shows the results of the structured approach applied to the main analysis for the subgroup of biopsy-naïve patients. Results are internally consistent, and consistent with the sources of evidence these drew upon. They mirror the high level of uncertainty in the evidence base.

The prevalence estimates inferred by the extended synthesis model are in line with those available in the literature (Section 4.7.1.1), perhaps closer to the lowest available estimate of cancer prevalence (i.e., low probability of no cancer). This is, however, expected, as the inferred prevalence in the extended model is bounded by a composite of all five tests, and is sampled from a distribution that allows for even higher cancer prevalences than those identified by the composite of all five tests.

Table 35 Distribution of test results, conditional accuracy and prevalence probabilities (mean and 95% CrI) according to ISUP grade for biopsy naïve individuals. Diagnostic accuracy extension to the evidence synthesis model. Results of main analysis

Network 1

		cognitive fusion					software fusion				
		(distribution of test results)					(distribution of test results)				
		0.516 [0.416,0.615]	0.186 [0.131,0.249]	0.136 [0.068,0.211]	0.098 [0.052,0.157]	0.064 [0.031,0.114]	0.457 [0.403,0.513]	0.173 [0.137,0.214]	0.196 [0.157,0.233]	0.108 [0.079,0.144]	0.066 [0.043,0.095]
		(accuracy matrix)					(accuracy matrix)				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.121 [0.007,0.238]	**No cancer**	1	0	0	0	0	1	0	0	0	0
0.318 [0.212,0.452]	**1**	0.829 [0.529,0.994]	0.171 [0.006,0.471]	0	0	0	0.671 [0.538,0.796]	0.329 [0.204,0.462]	0	0	0
0.262 [0.193,0.341]	**2**	0.300 [0.016,0.64]	0.362 [0.083,0.674]	0.338 [0.111,0.55]	0	0	0.251 [0.167,0.347]	0.204 [0.128,0.288]	0.544 [0.443,0.64]	0	0
0.183 [0.119,0.265]	**3**	0.189 [0.006,0.526]	0.140 [0.005,0.422]	0.192 [0.008,0.537]	0.479 [0.213,0.804]	0	0.224 [0.121,0.343]	0.059 [0.012,0.138]	0.207 [0.112,0.322]	0.510 [0.387,0.65]	0
0.116 [0.077,0.174]	**4 or 5**	0.125 [0.004,0.389]	0.111 [0.004,0.357]	0.111 [0.004,0.362]	0.101 [0.002,0.332]	0.552 [0.299,0.882]	0.111 [0.046,0.199]	0.047 [0.011,0.112]	0.130 [0.063,0.217]	0.140 [0.068,0.226]	0.573 [0.467,0.687]

Network 2

		(distribution of test results)					(distribution of test results)				
		0.460 [0.335,0.583]	0.250 [0.152,0.356]	0.127 [0.034,0.261]	0.131 [0.046,0.231]	0.033 [0.001,0.107]	0.348 [0.273,0.418]	0.223 [0.179,0.273]	0.232 [0.168,0.311]	0.115 [0.081,0.152]	0.082 [0.054,0.114]

		Combined cognitive fusion and systematic biopsy					**Combined software fusion and systematic biopsy**				
		(accuracy matrix)					**(**accuracy matrix)				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.121 [0.007,0.238]	**No cancer**	1	0	0	0	0	1	0	0	0	0
0.318 [0.212,0.452]	**1**	0.709 [0.289,0.987]	0.291 [0.013,0.711]	0	0	0	0.528 [0.206,0.824]	0.472 [0.176,0.794]	0	0	0
0.262 [0.193,0.341]	**2**	0.249 [0.01,0.689]	0.437 [0.07,0.836]	0.314 [0.028,0.78]	0	0	0.078 [0.001,0.273]	0.152 [0.011,0.384]	0.770 [0.523,0.975]	0	0
0.183 [0.119,0.265]	**3**	0.126 [0.002,0.488]	0.124 [0.003,0.449]	0.134 [0.002,0.482]	0.616 [0.148,0.981]	0	0.132 [0.005,0.441]	0.135 [0.005,0.411]	0.130 [0.004,0.403]	0.603 [0.338,0.92]	0
0.116 [0.077,0.174]	**4 or 5**	0.195 [0.004,0.618]	0.187 [0.005,0.603]	0.173 [0.004,0.561]	0.163 [0.004,0.543]	0.281 [0.006,0.865]	0.069 [0.001,0.282]	0.070 [0.001,0.265]	0.066 [0.001,0.27]	0.071 [0.001,0.266]	0.724 [0.402,0.98]

In terms of distribution of test results, the results obtained here (presented at the top of each accuracy matrix in Table 35) are consistent with those in the synthesis section (see Appendix 10 for detailed comparisons). In summary, within network 1 (which includes cognitive and software fusion), the results suggest that software fusion may retrieve a higher detection of cancer at ISUP grade 2 and above when compared to cognitive fusion, with the detection at ISUP Grade 2 being highest. These results are not statistically significant, in that credible intervals overlap significantly. Table 36 presents the information on distribution of test results converted onto detection rates at thresholds of categories. This information highlights that: software fusion presents a similar level of detection at ISUP grade 4-5, slightly increased detection of ISUP grade 3 or above of 1.3%, increased detection of ISUP grade 2 or above of 7.1% and increased detection at ISUP grade 1 or above of 5.9%.

Table 36: Proportion correctly identified and detection rates (mean and 95% CrI) with the different biopsy methods for biopsy naïve individuals. Diagnostic accuracy extension to the evidence synthesis model. Results of main analysis.

	Estimated detection rates with the different biopsy methods			
	Network 1		Network 2	
ISUP grade	**cognitive fusion**	**software fusion**	**combined cognitive fusion and systematic biopsy**	**combined software fusion and systematic biopsy**
4 or 5	0.064 [0.031,0.114]	0.066 [0.043,0.095]	0.033 [0.001,0.107]	0.082 [0.054,0.114]
3 to 5	0.162 [0.102,0.237]	0.175 [0.135,0.217]	0.164 [0.064,0.27]	0.197 [0.152,0.243]
2 to 5	0.299 [0.209,0.396]	0.370 [0.322,0.424]	0.290 [0.173,0.428]	0.429 [0.358,0.502]
1 to 5	0.484 [0.385,0.584]	0.543 [0.487,0.597]	0.540 [0.417,0.665]	0.652 [0.582,0.727]

In terms of accuracy, the results for network 1 suggest that software fusion is more accurate at detecting the correct category (the diagonal of the accuracy matrix is always higher for software fusion), with higher differences at ISUP grades 1 and 2.

The accuracy matrix results show that despite cognitive presenting a higher likelihood of an ISUP grade 1 result, there is an increased accuracy of software fusion at ISUP grade 1. This is due to, with cognitive fusion, individuals at higher ISUP categories being misclassified as grade 1. The accuracy matrix shows increased accuracy at ISUP grade 2 for software fusion, but retains a significant proportion inaccurately classified as ‘no cancer’ (with a probability of 0.25 95%CrI[0.17,0.34]) which is higher than the proportion innacurately classified as ISUP grade 1 (with a probability of 0.20 95%CrI[0.13,0.29]). A similar effect is observed in ISUP grade 3, where the likelihood of being classified as ‘no cancer’ is higher for software fusion than for cognitive fusion – probabilities 0.224 95%CrI [0.121,0.343] vs. 0.189 95%CrI [0.006,0.526]. This is a result of the increased detection at ISUP grade 2 not being matched by a similar level of detection at ISUP grade 1.

By multiplying the prevalence by the respective element of the accuracy matrix, the joint probability matrix is obtained. This matrix identifies, for a cohort with the mix of ISUP grades as per the prevalence estimates, the probability of both events, i.e. the probability of a particular 'true' ISUP grade and a particular test result. This matrix identifies that, at all grades, the probability of an accurate result is 0.524 95%CrI[0.411,0.628] for software fusion – 0.12 at no cancer, 0.10 at ISUP grade 1, 0.14 at ISUP grade 2, 0.09 at ISUP grade 3 amd 0.07 at ISUP grade 4 or 5. The probability of an accurate result is 0.413 95%CrI[0.256,0.583] for cognitive fusion – 0.12 at no cancer, 0.05 at ISUP grade 1, 0.09 at ISUP grade 2, 0.09 at ISUP grade 3 amd 0.06 at ISUP grade 4 or 5. The highest difference between software and cognitive is observed at ISUP grades 1 and 2 (approximately 5% increase in each with software). Notably, in terms of misclassification, the overall proportion of ISUP grade 3 identified as 'no cancer' is higher with software fusion 4.2% than with cognitive fusion (3.5%). This implies that the key trade-offs for software fusion are the benefits achieved by the general increase in detection, but particularly for ISUP grade 1 and 2, at the expense of a slightly higher proportion of grade 3's that will not be detected as cancerous.

Network 1

		(distribution of test results)					(distribution of test results)				
		0.516 [0.416,0.615]	0.186 [0.131,0.249]	0.136 [0.068,0.211]	0.098 [0.052,0.157]	0.064 [0.031,0.114]	0.457 [0.403,0.513]	0.173 [0.137,0.214]	0.196 [0.157,0.233]	0.108 [0.079,0.144]	0.066 [0.043,0.095]
		cognitive fusion					**software fusion**				
		(joint probability matrix)					(joint probability matrix)				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.121 [0.007,0.238]	**No cancer**	0.121 [0.007,0.238]	0	0	0	0	0.121 [0.007,0.238]	0	0	0	0
0.318 [0.212,0.452]	**1**	0.265 [0.136,0.413]	0.053 [0.002,0.141]	0	0	0	0.215 [0.122,0.335]	0.103 [0.063,0.149]	0	0	0
0.262 [0.193,0.341]	**2**	0.08 [0.004,0.184]	0.094 [0.022,0.178]	0.088 [0.026,0.145]	0	0	0.066 [0.038,0.104]	0.054 [0.03,0.084]	0.142 [0.102,0.185]	0	0
0.183 [0.119,0.265]	**3**	0.035 [0.001,0.108]	0.026 [0.001,0.081]	0.035 [0.002,0.105]	0.086 [0.032,0.146]	0	0.042 [0.016,0.079]	0.011 [0.002,0.029]	0.038 [0.017,0.071]	0.092 [0.062,0.129]	0
0.116 [0.077,0.174]	**4 or 5**	0.015 [0.000,0.048]	0.013 [0.000,0.044]	0.013 [0.000,0.043]	0.012 [0.000,0.04]	0.064 [0.031,0.114]	0.013 [0.005,0.026]	0.006 [0.001,0.014]	0.015 [0.006,0.031]	0.016 [0.007,0.032]	0.066 [0.043,0.095]

Network 2

		(distribution of test results)					(distribution of test results)				
		0.460 [0.335,0.583]	0.250 [0.152,0.356]	0.127 [0.034,0.261]	0.131 [0.046,0.231]	0.033 [0.001,0.107]	0.348 [0.273,0.418]	0.223 [0.179,0.273]	0.232 [0.168,0.311]	0.115 [0.081,0.152]	0.082 [0.054,0.114]
		Combined cognitive fusion and systematic biopsy					**Combined software fusion and systematic biopsy**				
		(joint probability matrix)					(joint probability matrix)				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.121 [0.007,0.238]	**No cancer**	0.121 [0.007,0.238]	0	0	0	0	0.121 [0.007,0.238]	0	0	0	0
0.318 [0.212,0.452]	**1**	0.227 [0.08,0.382]	0.091 [0.004,0.212]	0	0	0	0.171 [0.051,0.306]	0.147 [0.054,0.219]	0	0	0
0.262 [0.193,0.341]	**2**	0.066 [0.002,0.185]	0.114 [0.017,0.229]	0.082 [0.007,0.208]	0	0	0.021 [0.000,0.080]	0.041 [0.003,0.122]	0.199 [0.141,0.251]	0	0
0.183 [0.119,0.265]	**3**	0.023 [0.000,0.093]	0.022 [0,0.09]	0.025 [0.000,0.097]	0.112 [0.023,0.212]	0	0.025 [0.001,0.098]	0.026 [0.001,0.092]	0.025 [0.001,0.089]	0.106 [0.063,0.146]	0
0.116 [0.077,0.174]	**4 or 5**	0.023 [0.000,0.078]	0.022 [0.001,0.078]	0.02 [0.000,0.069]	0.019 [0.000,0.067]	0.033 [0.001,0.107]	0.009 [0.000,0.042]	0.009 [0.000,0.037]	0.008 [0.000,0.037]	0.009 [0.000,0.038]	0.082 [0.054,0.114]

Network 2 (including software and cognitive fusion combined with systematic biopsy) shows higher identification in the distribution of test results (due to the baseline used), but qualitative results are similar to those in network 1, noting that there is substantial uncertainty in these results. Detection rates at thresholds of categories show that cancer detection is expected to be higher with combined software, at all levels, but particularly at ISUP grade 2 or above where detection is 13.9% higher than with combined cognitive fusion and at ISUP grade 1 or above where detection is 9.2% higher that with combined cognitive fusion.

In terms of accuracy, at all grades, the probability of an accurate result is 0.655 95%CrI[0.471,0.816] for combined software fusion and 0.438 95%CrI[0.218,0.665] for combined cognitive fusion. For both combined strategies, the likelihood of a 'no cancer' result for ISUP grades 2 and 3 is still relatively high, but this is now comparable to the likelihood of an ISUP grade 1 result.

Subgroup analyses (previous negative biopsy)

In this analysis, the baseline distribution of test results for software fusion was sourced from Filson (2016)[96], but using the group of individuals recruited into this study that had previous negative biopsy results. However, the diagnostic accuracy evidence synthesis and the accuracy matrix are still sourced as per the main analysis, grounded on evidence over biopsy naïve patients. Table 37 presents summary results of distribution of test results and prevalence probabilities and results of the accuracy matrices are presented in Appendix 2.

The summary results in Table 37 illustrate that, for individuals with a previous negative biopsy, a significantly increased proportion of 'no cancer' results are expected in relation to biopsy naïve individuals. This impacts the (implicit) prevalence estimates: for those with previous negative biopsy, the probability of no cancer is 43% (95% CrI 26% to 53%), while for biopsy naïve it is 12% (95% CrI 0.7% to 24%). In comparing software with cognitive fusion biopsy strategies, across both networks, we observe similar probabilities of ISUP grade 1, 3 and 4 or 5 results, and a slightly higher probability of ISUP grade 2 results for software strategies. This differs from the results of the synthesis model for ISUP grade 3 only, where the probability under combined cognitive fusion was slightly higher than for combined fusion software (Table 10). The accuracy matrix estimates (reported in Appendix 10) are similar to those estimated for biopsy-naïve individuals (main analysis, Table 35).

Table 37: Distribution of test results and prevalence probabilities (mean and 95% CrI) according to ISUP Grade (D) for subgroup and sensitivity analysis. Diagnostic accuracy extension to the evidence synthesis model.

		Distribution of test results	
		Network 1	Network 2

Prevalence	ISUP	cognitive fusion	software fusion	Combined cognitive fusion and systematic biopsy	Combined software fusion and systematic biopsy
Subgroup analysis (previous negative biopsy)					
0.428 [0.259,0.529]	**No cancer**	0.703 [0.618,0.776]	0.661 [0.611,0.709]	0.659 [0.561,0.752]	0.583 [0.513,0.649]
0.224 [0.138,0.39]	**1**	0.105 [0.071,0.155]	0.107 [0.082,0.136]	0.157 [0.096,0.241]	0.155 [0.114,0.198]
0.132 [0.091,0.188]	**2**	0.077 [0.035,0.138]	0.107 [0.082,0.136]	0.067 [0.015,0.152]	0.120 [0.074,0.181]
0.131 [0.079,0.199]	**3**	0.068 [0.033,0.120]	0.080 [0.055,0.111]	0.091 [0.027,0.165]	0.083 [0.057,0.117]
0.085 [0.053,0.127]	**4 or 5**	0.046 [0.021,0.085]	0.049 [0.031,0.073]	0.027 [0.001,0.081]	0.058 [0.037,0.084]
Sensitivity analysis to baseline distribution for biopsy naïve (PAIREDCAP's baseline, Mortezavi's accuracy)					
0.031 [0.001,0.092]	**No cancer**	0.368 [0.248,0.473]	0.314 [0.271,0.362]	NA	NA
0.226 [0.163,0.319]	**1**	0.191 [0.140,0.256]	0.169 [0.137,0.207]	NA	NA
0.322 [0.222,0.42]	**2**	0.196 [0.101,0.306]	0.263 [0.218,0.308]	NA	NA
0.252 [0.154,0.37]	**3**	0.145 [0.079,0.228]	0.098 [0.064,0.140]	NA	NA
0.169 [0.104,0.254]	**4 or 5**	0.101 [0.052,0.176]	0.098 [0.064,0.140]	NA	NA
Sensitivity analysis to accuracy matrix for biopsy naïve (Filson's baseline, Zhou's accuracy)					
0.170 [0.023,0.280]	**No cancer**	0.525 [0.433,0.620]	0.450 [0.400,0.509]	NA	NA
0.279 [0.196,0.400]	**1**	0.190 [0.131,0.256]	0.175 [0.140,0.217]	NA	NA
0.300 [0.211,0.436]	**2**	0.122 [0.062,0.201]	0.189 [0.147,0.236]	NA	NA
0.155 [0.109,0.223]	**3**	0.098 [0.053,0.158]	0.112 [0.082,0.146]	NA	NA
0.095 [0.067,0.136]	**4 or 5**	0.450 [0.400,0.509]	0.075 [0.053,0.103]	NA	NA

Sensitivity analysis

- Sensitivity analyses changes the main sources of evidence of the main analyses (on biopsy naïve patients): a first sensitivity analysis uses an alternative baseline distribution of test results for software fusion (from PAIREDCAP (2019),[88] and a second analysis uses an alternative source for accuracy matrix evidence (from Zhou et al. (2018)).[108].

In both these analyses, results for the accuracy matrices could only be presented for the first network because of increased uncertainty.

The summary results in Table 37 for the first sensitivity analysis indicate that results are sensitive to the distribution of test results. The PAIREDCAP study distribution showed a higher proportion of 'no cancer' identified with software fusion (31% vs. 46% in the main analysis grounded on Filson, Table 35), identical in ISUP grade 1, and higher proportions across all remaining ISUP categories (26%, 16% and 10%, respectively for ISUP grade 2, 3 and 4 or 5, vs. 20%, 11% and 7% in the main

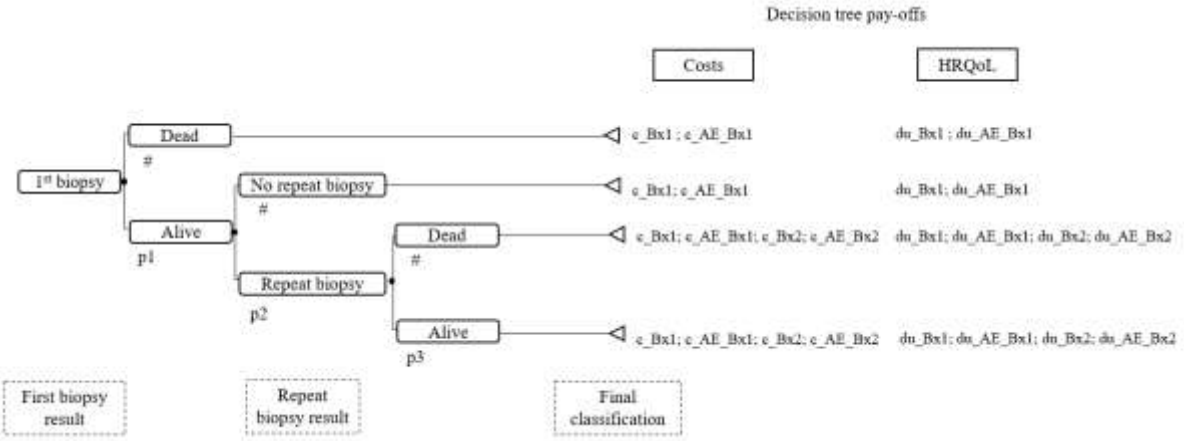
analysis grounded on Filson, Table 35). The distribution of test results for ISUP grades 4 or 5 are similar between software and cognitive fusion, but are significantly increased for software at ISUP grade 2, slightly increased at ISUP grade 3 and slightly reduced for ISUP grade 1.

The summary results in Table 37 for the second sensitivity analysis indicate that results on the distribution of test results are only slightly sensitive to the source of evidence on the accuracy matrix. Filson, Table 35). The main difference distribution of test results for ISUP grades 4 or 5 are slightly higher for software fusion in this analysis in relation to the main analysis in Table 35. The estimates of the accuracy matrices (shown in Appendix 2) show increased accuracy (in classifying individuals in the right category) for both technologies in relation to the main analysis in Table 35, which reflect the data from Zhou (2018).[108] However, differences between the technologies in the accuracy matrices are encountered in individuals with true ISUP grade 4 or 5 where the misclassified have an equal chance across being identified across all other categories in cognitive but are slightly less likely to be identified as no cancer or ISUP grade 1 with software fusion. For those in ISUP grade 2, sensitivity analysis indicates a low likelihood of the misclassified being identified as Grade 1 with software fusion (and therefore being more likely to be classified as 'no cancer'), which was not observed in the main analysis.

6.3.2 Diagnostic pathway

The diagnostic pathway is structured as a decision tree that captures adverse events, repeat biopsies and classifies individuals according to the result of the biopsy (or biopsies), and the true disease status (see Section 6.3.1), defined as ISUP grade for those with prostate cancer (ISUP grades 1, 2, 3, 4-5). Figure 8 shows a simplified schematic of the decision tree illustrating biopsy related mortality, sequence of biopsies, and cost and HRQoL pay-offs which apply for each strategy. The diagram does not show the biopsy related non-fatal events, as these do not modify the probability of moving forward in the diagnostic pathway. The probabilities of adverse events are applied as weights to adjust the branch costs and HRQoL pay-offs. The diagram also does not show how the classification is established conditional on the true disease state and test accuracy at each biopsy, or how the classification conditions the probability of repeat biopsy; this is illustrated subsequently in this subsection (Table 38).

Figure 8 Decision tree schematics



●, probability node;◁, terminal node; #, complement probability (1-probability); AE, adverse events; Bx, biopsy; c_, cost; du_, disutility; p1, probability of surviving the 1st biopsy; p2, probability of repeat biopsy (is conditional on 1st biopsy result); p3, probability of surviving 2nd biopsy.

All individuals who undergo the first biopsy are at risk of biopsy related non-fatal and fatal adverse events. The mortality risk corresponds to the complement of probability p1. For those who survive the first biopsy, the probability of receiving a repeat biopsy (p2) is conditional on the result of the first biopsy. Individuals who test positive at first biopsy (biopsy result ISUP grade ≥2) and survived the first biopsy receive no further testing (p2=0). Those who test negative (no prostate cancer or ISUP grade 1) and survived the first biopsy have a probability of undergoing repeat biopsy (p2), with the remaining individuals receiving no further testing. The individuals who receive a repeat biopsy are again exposed to biopsy mortality risk (1-p3), and to a probability of having non-fatal biopsy adverse events. Time is not modelled within the decision tree, so events are assumed to occur instantaneously (or in rapid sucession prior to long-term model entry); this is in line with the other cohort models examined in Section 5.6.

The decision tree models repeat biopsy for a proportion of individuals who have a negative first biopsy result. In the base-case, this proportion is not conditional on whether the strategy includes a cognitive or software fusion component. The base-case analysis assumes that the proportion of repeat biopsy is only conditional on the result of the first biopsy (15.45% and 5%, if the result of the first biopsy indicated a lesion with ISUP grade 1 and no prostate cancer, respectively) as per a previous DAR.[126]

Similar to a previous DAR,[126] we assume the same rates of biopsy complications per biopsy approach for the first and repeat biopsies. However, because we assume a different distribution between transperineal and transrectal biopsy for the first and repeat biopsies in the diagnostic pathway, the

repeat biopsy complication rates reflect a higher proportion of transperineal biopsy (10% GATP and 60% LATP) compared to first biopsy (65% LATP) (see Section 6.3.7.1).

In the base-case scenario, the diagnostic performance of the repeat biopsy is assumed the same as of the first biopsy. The model allows exploring a degradation in the diagnostic performance of repeat when compared to first biopsy; the impact of applying this alternative assumption is assessed through scenario analysis.

Table 38 illustrates the set of possible results (classification) of first, repeat biopsy (for the proportion of individuals who undergo a repeat biopsy), and final classification, according to the joint probabilities of being classified in ISUP grade j with test k conditional on being in true latent category i. The final classification is assumed to correspond to the highest result of the two biopsies, since we are assuming that misclassification at a higher category is not possible. Misclassification at the terminal nodes (final classification) of the model is highlighted in italic in Table 38.

Table 38 Test sequence and classification in the diagnostic pathway

True disease state	1st biopsy classification	Repeat biopsy	Repeat biopsy classification	Final classification
No PCa	No PCa	95%* No	-	No PCa
		5%* Yes	No PCa	No PCa
ISUP grade 1 (Gleason score 3+3)	No PCa	95%* No	-	*No PCa*
		5%* Yes	No PCa	*No PCa*
	ISUP grade 1	85%* No	-	ISUP grade 1
		15%* Yes	No PCa	ISUP grade 1
			ISUP grade 1	ISUP grade 1
ISUP grade 2 (Gleason score 3+4)	No PCa	95%*No	-	*No PCa*
		5%* Yes	No PCa	*No PCa*
	ISUP grade 1	85%* No	-	*ISUP grade 1*
		15%* Yes	No PCa	*ISUP grade 1*
			ISUP grade 1	*ISUP grade 1*
			ISUP grade 2	ISUP grade 2
	ISUP grade 2	No	-	ISUP grade 2
ISUP grade 3 (Gleason score 4+3)	No PCa	95%*No	-	*No PCa*
		5%* Yes	No PCa	*No PCa*
	ISUP grade 1	85%* No	-	*ISUP grade1*
		15%* Yes	No PCa	*ISUP grade1*
			ISUP grade 1	*ISUP grade1*
			ISUP grade 2	*ISUP grade2*
			ISUP grade 3	ISUP grade 3
	ISUP grade 2	No	-	*ISUP grade 2*
	ISUP grade 3	No	-	ISUP grade 3
ISUP grade 4-5 (Gleason score ≥8)	No PCa	95%*No	-	*No PCa*
		5%* Yes	No PCa	*No PCa*
	ISUP grade 1	85%* No	-	*ISUP grade 1*
		15%* Yes	No PCa	*ISUP grade 1*
			ISUP grade 1	*ISUP grade 1*

			ISUP grade 2	*ISUP grade 2*
			ISUP grade 3	*ISUP grade 3*
			ISUP grade 4-5	ISUP grade 4-5
	ISUP grade 2	No	-	*ISUP grade 2*
	ISUP grade 3	No	-	*ISUP grade 3*
	ISUP grade 4-5	No	-	ISUP grade 4-5

ISUP, International Society of Urological Pathology; PCa, prostate cancer

We note the (first and repeat) biopsy results are assigned in the decision tree immediately before the biopsy mortality risk is applied, meaning the proportion of individuals in each category is adjusted by the proportion who survived the biopsy procedure (assuming the same mortality risk applies to all individuals regardless of true disease category and biopsy result). Similarly, we assumed that the biopsy adverse events apply to all individuals who undergo a biopsy procedure.

The costs and QALY pay-offs in the decision tree capture the short-term impacts of first and repeat biopsy. First biopsy cost pay-offs apply to all branches and include the cost of the biopsy procedure and of associated adverse events. Similarly, the QALY pay-offs of the first biopsy also apply to all decision tree branches. These QALY pay-offs aim to quantify the QALY loss associated with biopsy procedural complications. The repeat biopsy related costs (including the same cost categories as for the first biopsy) and repeat biopsy complications QALY loss apply only to the decision tree branches which include a repeat biopsy.

The costs of the biopsy procedure vary by strategy to reflect the differences in cost between cognitive fusion and software fusion with each of the MRI fusion systems modelled (see Section 6.3.7.1 for the estimation of biopsy procedure costs). The biopsy procedure and adverse events costs are both specific to the biopsy approach (LATP, GATP or LATRUS); these costs are estimated as a weighted average of the costs by biopsy approach (where the weights correspond to the proportion of LATP, GATP and LATRUS for each biopsy in the strategy). The QALY loss from biopsy related complications also varies by biopsy approach to reflect the different biopsy complication rates by biopsy route of access (transperineal or transrectal) and, therefore, is also estimated as a weighted average by biopsy approach.

6.3.2.1 Clinical management conditional on biopsy final classification

There are 15 possible final classifications at the end of the diagnostic pathway, which are as follow:

1. For individuals correctly classified:
    1.1. Diagnosed as having no prostate cancer and without prostate cancer;
    1.2. Diagnosed as ISUP grade 1 and with ISUP grade 1;
    1.3. Diagnosed as ISUP grade 2 and with ISUP grade 2;
    1.4. Diagnosed as ISUP grade 3 and with ISUP grade 3;

1.5. Diagnosed as ISUP grade 4-5 and with ISUP grade 4-5;

2. For individuals misclassified:
   2.1. Diagnosed as having no prostate cancer and with:
      2.1.1. ISUP grade 1;
      2.1.2. ISUP grade 2;
      2.1.3. ISUP grade 3;
      2.1.4. ISUP grade 4-5;
   2.2. Diagnosed as ISUP grade 1 and with;
      2.2.1.ISUP grade 2;
      2.2.2. ISUP grade 3;
      2.2.3. ISUP grade 4-5;
   2.3. Diagnosed as ISUP grade 2 and with;
      2.3.1.ISUP grade 3;
      2.3.2. ISUP grade 4-5;
   2.4. Diagnosed as ISUP grade 3 and with;
      2.4.1.ISUP grade 4-5.

The clinical management for each of these possible classifications is dependent on the diagnosed category. As detailed in Section 2.2, current clinical guidance[10] recommends that individuals diagnosed as having localised or locally advanced disease are involved in decisions about the management of their disease, with the range of management options offered varying as a function of their prognostic risk. Thus, patients with lower CPG scores (better prognosis) are offered more conservative management (active surveillance) with option to undergo radical treatment, while those with higher CPG scores are offered radical treatment as the preferred management option.

The diagnostic performance evidence only allows classifying patients according to their histopathological information (i.e., ISUP grade), which is only part of the prognostic information used to determine the CPG scores. Therefore, we made a simplifying assumption that ISUP grade can be used as a proxy for the individuals' CPG score (e.g., CPG1 = ISUP grade 1), to allow establishing the evidence linkage between classification and clinical management and subsequently from this to treatment outcomes. Henceforth, we refer to the classification in the model in terms of CPG score, assuming interchangeability between ISUP grades and CPG scores. The treatment options for localised and locally advanced disease include active surveillance or radical treatment. Radical treatment includes radiotherapy (consisting in the model of brachytherapy or external beam radiotherapy for costing purposes [see Section 6.3.10]), and radical prostatectomy.

For individuals identified as having prostate cancer, the model assigns varying proportions of active surveillance and radical treatment, according to diagnosed CPG score (See Section 6.3.4.1). All

patients in the localised disease health states receive monitoring, with the set of monitoring tests and schedule varying according to whether they are receiving active surveillance or radical treatment. Individuals without a prostate cancer diagnosis also receive monitoring, but its regime is less intensive compared to individuals diagnosed with prostate cancer and is time limited (maximum of 10 years).

Prostate cancer treatment is associated with adverse events, such as sexual, urinary and bowel dysfunction, with rates of adverse events varying by treatment (see Section 6.3.5.2). Adverse events from prostate cancer management are associated with disutility and costs of managing these events, which are quantified within the long-term model.

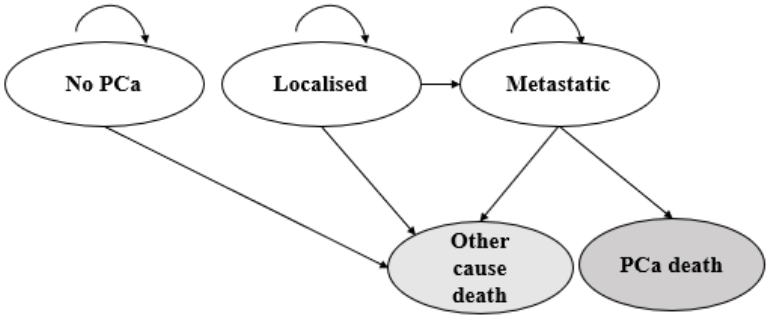
6.3.3 Modelling of long-term outcomes

6.3.3.1 Overview of the decision analytic model

The long-term outcomes of the model cohort conditional on latent true disease status, the diagnosed disease category and prostate cancer management assigned, are quantified in a state transition Markov model. The model has yearly cycles (with a half-cycle correction applied) and a lifetime time horizon (40 years).

The core structure of the model is illustrated in Figure 9. Individuals who survived the biopsy procedure(s) in the diagnostic pathway can enter the model through the no prostate cancer state if they are disease free or the localised (and locally advanced) disease state if they have prostate cancer. Patients with prostate cancer at model entry can remain in the localised disease health state or transition to the metastatic disease state at each yearly model cycle. The individuals who died due to the diagnostic procedure enter the 'other cause' death state, one of the two absorbent states in the model (highlighted in grey in Figure 9). Transitions to the other cause death state are possible from the 'no prostate cancer', localised and metastatic disease health states, with the same probability as the general population (see Section 6.3.3.7). The only other possible transition for the 'no prostate cancer' state is to remain in the same state (i.e., the model does not consider that individuals can develop prostate cancer, so disease progression is not modelled for those who do not have the disease at model entry). The metastatic health state is modelled as three tunnel health states (not illustrated in this diagram, see Section 6.3.11, Figure 11), where individuals can only stay in the two first tunnels states for a maximum of one year. Patients who transition to the metastatic health state can only remain in that health state or die. Prostate cancer mortality only applies to patients in the metastatic disease states.

Figure 9 Long-term outcomes Markov model structure



Localised disease health states

True disease status	**No PCa**	CPG 1	**CPG 1**	CPG 2	CPG 2	**CPG 2**	CPG 3	CPG 3	CPG 3	**CPG 3**	CPG 4-5	CPG 4-5	CPG 4-5	CPG 4-5	**CPG 4-5**
Diagnosed	**No PCa**	No PCa	**CPG 1**	No PCa	CPG 1	**CPG 2**	No PCa	CPG 1	CPG 2	**CPG 3**	No PCa	CPG 1	CPG 2	CPG 3	**CPG 4-5**

CPG, Cambridge Prognostic Group; PCa, prostate cancer

There are 15 possible localised disease health states (illustrated in the box below the model schematics), each reflecting the final classification (here expressed as CPG scores) attributed by the diagnostic pathway and the different treatments assigned conditional on the diagnosed category in the final classification.

Over the next subsections we provide details on the parameterisation of long-term transition probabilities.

6.3.3.2 Inference sub-model (disease progression by CPG and treatment intensity)

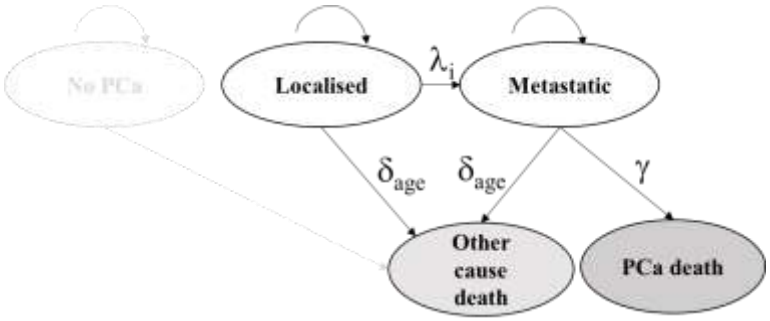
The decision analytic prostate cancer model requires consideration of the impact of treatment decisions according to diagnostic accuracy. Treatment decisions are currently grounded on the identification of CPG groups, and therefore the outcomes component of the model aims to reflect: 1) differences in outcomes across the CPG risk groups that underlie treatment decisions in clinical practice, and 2) the impact of different treatment intensities on each of these risk groups. Our conceptualisation review has not identified any previous cost-effectiveness model where treatment outcomes for 5-level CPG groups have been considered (see Section 5.6). Therefore, an estimation strategy was developed in this assessment grounded on the targeted review of evidence on the long-term outcomes of prostate cancer (see Section 4.7.2).

The brief overview of the wider literature highlights that, whilst there is evidence on the effectiveness of radical vs. 'conservative' treatment options in delaying progression to metastatic disease, there are limited mortality benefits observed within the follow-up of clinical trials in this area. Also, we did not

find evidence on treatment effectiveness stratified by CPG scores, despite the prognostic ability of the 5-level score for prostate cancer specific death having been demonstrated in a large UK-based observational study.[122]

The aim of the inference model is therefore to pull existing evidence together to predict differences in progression to metastatic disease by 5-level CPG score and by treatment. Given this has not been directly observed, a calibration model was developed to infer these. The calibration model uses the structure of the decision analytic model in Figure 9, but without considering the 'no prostate cancer' health state, which has, thus, been faded out in the diagram.

Figure 10 Calibration model structure and parameters



The model structure is underpinned by the following assumptions. All individuals are assumed to begin with localised disease. They can continue to have localised disease, progress to metastatic disease or die from causes other than prostate cancer. The speed of progression to metastatic disease is expected to depend on CPG group and is given by λ_i, where the index *i* reflects the CPG group. Other cause mortality is age-specific and is determined by δ_{age}. Those with metastatic disease may 1) continue to live with metastatic disease, 2) die from prostate cancer or 3) die from other causes. Following the NICE NG 131 model[133], it was assumed that death from prostate cancer could only occur after metastatic disease. The model was parameterised for each CPG score of interest to this assessment (CPG 1, 2, 3, and 4 and 5 combined).

The inference procedure is undertaken in two parts.

6.3.3.3 Part 1: Identifying rates of progression to metastatic disease by CPG, λ_i

This part uses calibration. For any calibration process, two sets of parameters are of interest. The first concerns model parameters, some of which are unobserved and the target of inference, and others are observed and therefore evidence directly informs these. The second set concerns calibration targets, which are functions of the model parameters that have been observed and are used to identify the unobserved parameters under the model structure and other observed inputs. Table 39 lists the

calibration parameters and targets and presents the results of the calibration model. A more detailed description of these parameters and their evidence sources is presented below in turn.

Table 39 Calibration model parameters and targets, and calibration results

	Description	Source	Parameter value	Results
PART 1	**Calibration targets**			
	10-year PC death by CPG group at diagnosis	Gnanapragasam[122], pooled results for testing and training sets	10-year PC survival (SE) [a, b parameters of a Beta distribution]: G1: 0.968 (0.007) [586, 19] G2: 0.938 (0.010) [577, 38] G3: 0.871 (0.016) [356, 53] G4/5: 0.763 (0.052) [50, 16]	-
	Calibration model parameters			
	Unobserved rate of progression from localised to metastatic disease, by CPG: λ_1, λ_2, λ_3, λ_4	Unobserved (calibration parameters)	NA	Rate (SE): G1: 0.0101 (0.00236) G2: 0.0229 (0.00403) G3: 0.0645 (0.01058) G4/5: 0.1788 (0.08641)
	Observed rate of PC mortality from metastatic disease, γ	STAMPEDE[151]	yearly rate of PC mortality in ADT arm of 0.162 (SE 0.0073), calculated from 5-years PC mortality	-
	Observed rate of death from other causes, age-specific	ONS life tables (2000-2002)[152]	Assumed mean age for each CPG group G1: 66.2 G2: 68.14 G3: 71.13 G4/5: 72.18	-
PART 2	Proportions under radical treatment vs. conservative management, by CPG	Gnanapragasam[122]	G1: 0.53 G2: 0.70 G3: 0.81 G4/5: 0.95	-
	Rate ratio for the development of metastasis of radical vs. conservative treatment	ProtecT[55]	Rate ratio = 0.43 95% CI[0.26 to 0.72], log rate ratio mean = -0.834, SE = 0.2545	-
	Rate of progression from localised to metastatic disease, by CPG and by treatment	Unobserved	NA	Conservative management $\lambda_1^{(0)}$= 0.0143 (0.00357) $\lambda_2^{(0)}$= 0.0380 (0.00832) $\lambda_3^{(0)}$= 0.1197 (0.02812) $\lambda_4^{(0)}$= 0.3950 (0.22287) Radical treatment: $\lambda_1^{(1)}$= 0.0063 (0.00184) $\lambda_2^{(1)}$= 0.0165 (0.00357) $\lambda_3^{(1)}$= 0.0516 (0.00964) $\lambda_4^{(1)}$= 0.1674 (0.08066)

ADT, Androgen deprivation therapy; CI, confidence interval; CPG, Cambridge Prognostic Group; NA, not applicable; PC, prostate cancer; SE, standard error

Calibration targets

Our calibration target is 10-year prostate cancer specific mortality according to CPG group at diagnosis of localised disease, as reported in Gnanapragasam et al. (2016).[122] Our analysis combines groups 4 and 5, as the recommended treatment is the same for both groups.[10] We used a single data point for each CPG group of interest, at 10-year follow-up, in the calibration model. We used WebPlotDigitizer [153] to extract point estimates and upper and lower confidence intervals for prostate cancer survival at 10 years (3652 days) from both training and validation sets in Gnanapragasam et al. (2016)[122] (from Figures 1A and 2A). Standard errors were calculated by considering the average distance between the point estimate and the upper and lower confidence limits (where both were available). The figures were then combined across datasets to derive a single estimate for each CPG, by weighting according to the inverse of their precision (analogously to a fixed-effect meta-analysis). Values for the combined CPG 4 and 5 group were derived by pooling the distributions, i.e., assuming that the variance of the combined group is the weighted sum of the variances in each group. To describe the survival probabilities probabilistically, the parameters of Beta distributions were specified using the method of moments. The estimates of 10-year prostate cancer survival and the parameters of the Beta distributions used to describe this in the calibration model are presented in Table 39. When simulating from the Beta distribution to run the calibration, we preserved the ordering ensuring survival is highest in group 1 then group 2, group 3 and group 4-5.

Calibration model parameters

The rates of progression from localised to metastatic disease by CPG score (λ_i, where *i* represent the CPG score groups of interest) were the unobserved parameters we sought to achieve inference on.

The remaining model parameters were observed. Prostate cancer specific mortality was assumed to only be possible after progression to metastatic disease, and therefore to inform this model parameter we used outcomes reported from STAMPEDE, a UK study.[151] Data from STAMPEDE's control arm were used, as long-term hormonal treatment was the standard of care at the time the study informing the calibration target was conducted. The individual patient data was reconstructed from the published Kaplan Meier curve using the Guyot algorithm[154] and a Weibull distribution was fitted using the *flexsurv* package in R.[155] Prostate cancer survival at 5 years predicted by the fitted Weibull function was 40.6% (95% confidence interval from 43.9 to 47.0%), which was converted onto a rate assuming constant hazard. This resulted in a mean hazard of 0.162 and a 95% confidence interval from 0.1777 to 0.1492. Assuming a symmetrical distribution this implies a standard error of 0.0073.

ONS life tables for men[152] were used to parameterise the transitions to death from other causes (both from localised disease and from metastatic disease). Life tables were used for the years 2000-2002 to approximate the mortality at the time of the Gnanapragasam study (2000 – 2010). The average age at the start of that study differs by risk group according to data reported in the NICE model.[133] Using

linear interpolation, we extended the 3 risk groups reported in the NICE report to the 4 risk groups we are considering. The ages assumed were: 66.2 years for Group 1, 68.14 years for Group 2, 71.13 years for Group 3 and 72.18 years for Groups 4-5. Due to the large sample size underlying the life tables, we did not consider this parameter uncertain.

Analysis methods

Using the parameters and targets described above we ran the calibration analysis in the software package R, according to the algorithm below:

> (1) Sample a value from the uncertainty distribution for the target (prostate cancer mortality at 10 years, for each risk group) and the known model inputs (metastatic mortality rate).
> (2) For each risk group, identify the value of the rate of transition from localised to metastatic disease (λ_i) that is consistent with the prostate cancer mortality at 10 years sampled in step 1. Record the result.
> (3) Repeat steps 1 and 2, 10,000 times.

The *optim* function in R was used for the second step in this algorithm.[36] To find the rate consistent with the target we defined a discrete time Markov model with the structure in Figure 10, and determined that its predicted 10-year survival should be compared against the target value. The loss function used was the squared distance from the proposed value to the target value. The Brent method was used with lower and upper bounds of 0 and 10 respectively.[156]

Results

The results from the calibration procedure regarding the unobserved rate of progression from localised to metastatic disease by CPG are shown in Table 39.

6.3.3.4 Part 2: Identifying the effect of treatment on the rates of metastasis

The estimated rate of progression to metastatic disease from the calibration exercise above reflect outcomes with current practice, which comprises of a mix of radical and conservative treatment. In part 2, we back-calculate how these rates differ for the proportions treated with radical and conservative treatment observed in Gnanapragasam et al. (2016)[122] using an external estimate of effect for radical treatment.

Gnanapragasam et al. (2016)[122] reported the treatment mix by risk group observed in UK clinical practice during the years 2000-2010. The treatment categories considered were: conservative management, brachytherapy, primary ADT (PADT), radical prostatectomy and radical radiotherapy. We further grouped treatments into two categories: conservative management and all other options, which we considered "radical treatment" (see appendix for further details and comparison to NICE guidance). The split by risk group is shown in Table 39.

The rates of progression to metastatic disease inferred in the calibration step (part 1) reflect the treatment allocations in Gnanapragasam et al. (2016)[122] (see Table 39). To consider the rates of progression with and without radical treatment, we disentangle the effect of treatment by considering that the pooled estimate of the rate of progression is a weighted average of the rates under radical treatment and conservative management (weighted by the proportions treated). The rates under radical treatment are assumed to be the rates under conservative treatment multiplied by a rate ratio sourced from external evidence. For such, we use the treatment effect from ProtecT,[55] the most recent UK study identified in the targeted review of the literature (see Section 4.7.2). The rate ratio data for radical treatment pooled the ProtecT results for radical prostatectomy and radiotherapy, retrieving an estimate of 0.43 (95% confidence interval from 0.26 to 0.72). Note that this estimate is similar to the USA based PIVOT study which estimated a HR for developing bone metastasis for radical prostatectomy of 0.40 (0.22 to 0.70).

Results

The results from the second part of the inference model regarding the rate of progression from localised to metastatic disease by CPG and by treatment are shown in Table 39.

6.3.3.5 Comparison of estimates with those from recent UK cost-effectiveness models

The review in Section 5.6 identified that cost-effectiveness models typically consider the increased, or earlier, identification of prostate cancer cases to affect health outcomes by modifying the likelihood of progression to metastatic disease (via earlier, or more appropriate, cancer treatment). Two approaches are used in the long-term outcome component of these models. The first way is to condition speed of progression on true risk group at the time of (or close to) model start. Models that use this approach typically focus on the diagnostic pathway (leading on to treatment decisions). Implicitly, future changes in disease status or in further treatment are implicitly considered in the evidence informing the likelihood of progression over time. One such model is the PROMIS long term model.[135]

The second approach implemented is to model explicitly progression across risk groups over time spent in model. Such explicit modelling of progression allows more granularity in the evaluation of monitoring, observation or watchful waiting type strategies, which in turn will determine future treatment decisions. Of the UK cost-effectiveness models, the long term inference model developed to inform the NICE guidelines model[133] (also used in the Southampton DAR[126]) is of such a kind.

We compared predictions of prostate cancer specific mortality at 2, 5, 10 years by risk group and treatment from our inference model with those of other UK relevant models: the NICE guidelines model[133] (used in the Southampton DAR[126]) and the PROMIS model[135, 136].

In summary, the outcomes component of the PROMIS model[135, 136] calibrated the probability of progressing to metastatic disease by risk category and treatment received. Calibration targets were survival data and proportion of patients with metastases by treatment arm from PIVOT[119] (risk stratified) considering mortality in the metastatic subgroup from the STAMPEDE trial.[123]

The NICE guidelines model also used calibration to derive transitions between risk groups and to metastatic disease (over time), under the assumption that patients would have to be high risk before developing metastatic disease.[133] The calibration targets (risk stratified) were, for undiagnosed cancers, metastases risk from the watchful waiting arm in SPCG4[145] and, for diagnosed cancers, cancer-specific survival from Gnanapragasam (2016)[122]. For both groups, mortality in the metastatic subgroup from the STAMPEDE trial[59] was considered.

To compare the different models, results were conditioned on risk group and treatment. To condition on risk group, across all model prevalence was set to 100% for each of the risk groups in turn. To condition on treatment, diagnostic accuracy was either set at 100% (to secure all patients are diagnosed and treated) or at 0% (to reflect the values if all patients are undiagnosed and untreated). Where relevant, diagnosis due to symptom presentation was not allowed. Where relevant, treatment allocation was set to 100% conservative management or, alternatively, to 100% radical treatment. To derive prostate cancer specific mortality in the PROMIS model, we only considered mortality in individuals with metastatic disease, and subtracted general mortality. The results of these analyses are presented in Table 40.

Table 40: Comparison of 2, 5 and 10-year PCa mortality predictions between alternative long-term outcome models

True disease status	Final classification/ treatment	PCa mortality at …		
		2 years	5 years	10 years
NICE guidelines model[133]				
LR	No PCa	<0.1%	0.7%	7.1%
	LR	<0.1%	0.1%	1.6%
IR	No PCa/LR	0.3%	3.9%	19.3%
	IR	<0.1%	0.7%	5.2%
HR	No PCa/LR	1.6%	9.5%	28.9%
	IR/HR	0.6%	4.0%	14.7%
PROMIS[135, 136]				
LR	WW	0.1%	0.9%	3.0%
IR	WW	0.3%	2.1%	6.4%
	RP	0.1%	0.8%	2.6%
HR	WW	0.3%	2.6%	7.8%
	RP	0.1%	0.9%	2.9%
De novo inference model by treatment				
CPG1	conservative	0.2%	1.5%	4.4%

	radical	0.1%	0.6%	2.0%
CPG2	conservative	0.5%	3.5%	9.8%
	radical	0.2%	1.6%	4.6%
CPG3	conservative	1.3%	8.6%	20.7%
	radical	0.6%	4.1%	10.8%
CPG4 or 5	conservative	3.5%	18.9%	36.2%
	radical	1.7%	10.7%	24.6%
De novo inference model with weighted treatment estimates				
CPG1	weighted	0.1%	0.4%	3.1%
CPG2	weighted	0.3%	2.2%	6.2%
CPG3	weighted	0.7%	5.0%	13.0%
CPG4 or 5	weighted	1.8%	11.2%	25.6%

CPG, Cambridge Prognostic Group; HR, high-risk; LR, low-risk; IR, intermediate-risk; PCa, prostate cancer; RP, radical prostatectomy; WW, watchful waiting

This table highlights that there are are marked differences between the predictions, which are primarily due to the sources of long-term outcome evidence these inference models relied upon, which differed. Table 41 depicts PCa mortality at 2, 5 and 10 years observed within the studies that served as calibration targets for the different models.

Table 41: PCa mortality at 2, 5 and 10 years observed within the studies that served as calibration targets for alternative long-term outcome models

Study	Population + treatment		PCa mortality at …		
			2 years	5 years	10 years
PIVOT[119]	LR	Obs	5.7%	13.8%	35.1%
		RP	3.4%	10.4%	29.6%
	IR	Obs	8.1%	16.3%	43.4%
		RP	3.3%	10.0%	31.4%
	HR	Obs	8.8%	24.8%	46.0%
		RP	3.8%	22.8%	51.5%
SPCG4[145]	LR	WW	<0.1%	<0.1%	4.5%
		RP	<0.1%	0.9%	3.3%
	IR	WW	<0.1%	3.8%	17.8%
		RP	<0.1%	1.9%	7.8%
	HR	WW	1.9%	7.1%	22.6%
		RP	<0.1%	3.9%	16.9%
Ganapragasam (2016)[122], 3-tier risk group*	LR	As per clinical practice	<0.1%	0.1%	3.1%
	IR		0.1%	2.0%	8.6%
	HR		2.1%	10.0%	23.4%
Ganapragasam (2016)[122], 5-tier risk group*	CPG1	As per clinical practice	<0.1%	1.0%	4.2%
	CPG2		<0.1%	1.7%	7.0%
	CPG3		<0.1%	3.5%	13.2%

	CPG4		0.1%	5.3%	17.7%
	CPG5		5.7%	19.4%	38.1%
	CPG45**		2.1%	10.2%	24.8%

*weighted average between training and testing datasets; **weighted average between CPG4 and CPG5; CPG, Cambridge Prognostic Group; HR, high-risk; LR, low-risk; IR, intermediate-risk; Obs, observation; PCa, prostate cancer; RP, radical prostatectomy; WW, watchful waiting

6.3.3.6 Parameterisation of the prostate cancer health states transition probabilities

The transition probabilities from each of the 15 localised disease health states to metastatic disease were informed by calibration as described above (see Section 6.3.3.2). The calibration estimated the transition rates by true disease status and treatment assigned (active surveillance or radical treatment). Transition probabilities were subsequently estimated by weighting the annual transition rates according to the treatments assigned based on diagnosed category (see Section 6.3.4.1, and then converted to annual transition probabilities assuming constant hazards over time (i.e., an exponential time to event distribution).

For patients in the metastatic disease health state transitions to prostate cancer death were informed by prostate cancer specific death from the UK STAMPEDE trial.[151] As described previously, a Weibull distribution was fitted using the *flexsurv* package in R data to the reconstructed individual level prostate cancer mortality data for the standard of care (ADT) arm (metastatic patient subgroup) in Clarke et al. (2019)[151] The choice of parametric distribution was in line with a recent NICE technology appraisal (TA) evaluating enzalutamide in combination with ADT for hormone sensitive metastatic cancer and based on a visual fit assessment conducted by the EAG (a full assessment of survival curve fit was considered out of scope for this assessment, so a targeted approach was taken). This baseline probability was parametrised in the executable model based on the *flexsurv* estimated Weibull coefficients (with a multivariate normal distribution fitted using the corresponding Cholesky decomposition for the PSA) and then adjusted by the effectiveness of contemporaneous combination treatments (i.e., in addition to ADT) weighted HR according to the current treatment distribution (see Section 6.3.4.1). The weighted hazard ratios for three combinations used in current clinical practice for first line metastatic prostate cancer compared to ADT alone were applied to the baseline probability of prostate cancer death to derive the metastatic to prostate cancer death transition probability. The combination treatments considered in the model included docetaxel (HR vs. ADT 0.78, 95% CI: 0.66 to 0.93)[151], enzalutamide (HR vs. ADT 0.66, 95% CI :0.53 to 0.81)[157] and apalutamide (HR vs. ADT 0.65, 95% CI: 0.53 to 0.79).[158] Lognormal distributions were fitted to each HR in the probabilistic model setup.

6.3.3.7 Other cause mortality

Age-dependent other cause mortality rates for men from ONS lifetables (2018-2020 collection period)[152] was used to estimate other cause death probabilities in the long-term model. Similar to the calibration model, parameter uncertainty was not considered for these inputs, due to the large sample size of the source dataset.

6.3.4 Treatment of prostate cancer

In Section 2.2.4, we stated that the clinical management (choice component) of individuals with a localised and locally advanced prostate cancer diagnosis in the model was conditional on i) diagnosed CPG score for the treatment component and ii) on the type of prostate cancer treatment (active surveillance or radical treatment) received for the routine monitoring component. For those in the metastatic health state, treatment includes androgen deprivation alone or in combination with other treatments. Here we present further details on the treatment distribution inputs in the model for i) localised and locally advanced disease and ii) metastatic disease.

6.3.4.1 Treatment of localised and locally advanced prostate cancer

NICE NG131 makes separate treatment recommendations by CPG score and conditional on patient preference and/or suitability for radical treatment (see Section 2.2.4) for individuals diagnosed with localised and locally advanced prostate cancer. In order to reflect treatment allocation based on the diagnosed CPG score and the patient level factors, we have sourced treatment allocation from Parry et al. (2020),[159] a study on the differences in localised and locally advanced treatment according to CPG in clinical practice in England. Our approach parameterising the treatment distribution contrasts to the approach taken in a previous DAR.[126] First, in the York model the distribution of active surveillance and radical treatments is conditional on the diagnosed disease status, whereas the Southampton DAR model conditioned this distribution on the "true" disease category. Second, this previous model sourced the treatment distribution mostly from Gnanapragasam (2016)[122] with further assumptions imposed on this distribution based on NPCA data. Both Parry et al. (2020)[159] and Gnanapragasam et al. (2016)[122] reported treatment distribution by CPG for cohorts of newly diagnosed with non-metastatic cancer. However, we preferred to source the treatment distribution from Parry et al. (2020)[159] to Gnanapragasam (2016),[122] because the data collection period is more recent (2014-2017 vs. 2000-2010) and had a higher sample size (61,999 vs. 10,139). Furthermore, Parry et al. (2020)[159] collected evidence from England, whereas Gnanapragasam et al. (2016)[122] was limited to data collected within the East of England Cancer Network area. We therefore considered the evidence in Parry et al. (2020)[159] study to be more contemporaneous and likely to be more reflective of current clinical practice than Gnanapragasam et al. (2016).[122]

Table 42 contrasts the prostate cancer management options distribution in the current and previous DAR. We note that the estimates applied in the York model for individuals diagnosed with CPG 2 and 3 suggest less use of radiotherapy and more use of radical prostatectomy compared to what was applied to individuals with intermediate risk disease in the Southampton DAR model.[126] There are also differences between studies in the proportion of individuals receiving active surveillance and watchful waiting.

Table 42 Localised and locally advanced disease treatment distribution

	Southampton DAR model[126]			**York model**			
Treatment choice based on	**"True" disease status**			**Diagnosed disease status**			
	Low-risk*	**Intermediate-risk***	**High-risk***	**CPG1**	**CPG2**	**CPG3**	**CPG4-5**
Active surveillance	95%	12.7%	0%	88.7%	51.6%	33.7%	24.1%
Radical prostatectomy	2%	21.9%	17.6%	6.6%	27.2%	26.3%	22.8%
Radiotherapy	3%	52.8%	52.4%	4.7%	21.3%	40.0%	53.1%
External radiotherapy	2.3%	48.7%	52.5%	3.6%	19.0%	38.2%	52.3%
Brachytherapy	0.7%	4.1%	0.9%	1.1%	2.3%	1.8%	0.8%
Watchful waiting	0%	12.7%	29%	0%	0%	0%	0%

*Low-risk assumed to correspond to CPG 1, intermediate-risk to CPG 2 and 3 (grey highlight), and high-risk to CPG 4 and 5

In the York model, we assumed that individuals would not be treated with watchful waiting, because this is a monitoring strategy for individuals for whom potentially curative treatment is not suitable (or do not wish to undergo this type of treatments). mpMRI to inform prostate biopsy decisions is currently only recommended for people who can undergo radical treatment,[10] so the exclusion of this treatment option was considered clinically plausible. We, therefore, assumed that individuals who were not treated with radical treatment, underwent active surveillance.

Parry et al. (2020)[159] did not report the proportion of individuals who were treated with brachytherapy, a radiotherapy that is more costly than external therapy. We assumed that the proportion of individuals treated with radiotherapy who underwent brachytherapy by CPG was the same as in Gnanapragasam et al. (2016),[122] with the remaining patients receiving external therapy. Furthermore, we assumed that all patients treated with radiotherapy also received ADT (length of treatment conditional on diagnosed CPG score), as per the Southampton DAR.[126] We note that current clinical guidance recommends 6 months of ADT before, during or after radiotherapy for individuals with CPG 2 to 5, and for treatment to continue for up to 3 years for people with CPG 4 and 5.

There is also an important structural difference in the choice component of the York model compared to the Southampton DAR.[126] As stated in Sections 6.3.3.2 and 6.3.3.5, the York model has flexibility

to reflect different treatment distributions between conservative (active surveillance) and radical treatment on disease progression, as the calibration model estimates disease progression rates by type of treatment and the derived transition probabilities for each localised disease health state are adjusted as a function of the treatment distribution per diagnosed CPG. In contrast, the calibrated disease progression probabilities in the Southampton DAR model[126] reflect the treatments received by the individuals in the outcome data used to derive them (i.e., Gnanapragasam (2016)[122] for the diagnosed states and Bill-Axelson (2014)[145] [observation arm]) and changes in the parameterisation of the treatment distribution only change how the cost and disutility of localised disease management are weighed in the model.

We fitted a Dirichlet probability distribution to the disaggregated observed count data by treatment type in Parry et al., (2020)[159] in the probabilistic parameterisation of the model.

6.3.4.2 Treatment of metastatic prostate cancer

Metastatic disease is treated initially with ADT alone or in combination, while disease is hormone-sensitive. Once disease progresses to hormone resistance, ADT is stopped and individuals will receive subsequent treatments.

Initial metastatic disease treatment (hormone sensitive metastatic cancer) was assumed to consist of a mix of ADT alone and in combination with docetaxel, enzalutamide and apalutamide, similarly to a previous DAR.[126] We updated the distribution of metastatic treatments in the Southampton DAR to reflect the 74% reduction in the use of docetaxel between 2019 and 2020 suggested by the NPCA 2021 report.[4] Therefore, in the York model we assumed that 9% of individuals with hormone sensitive metastatic cancer would be treated with docetaxel in combination with ADT (in contrast with the 36% in the Southampton DAR). We assumed that the difference in the proportion of treated with docetaxel between the two models (27%) would receive enzalutamide instead, since the NPCA 2021 report[4] also suggested a considerable increase on the use of this alternative treatment. We have sourced the proportion of treatment with ADT alone and in combination with apalutamide directly from the Southampton DAR.[126] The metastatic treatment distribution applied in the two models is reported in Table 94 (Appendix 11).

As mentioned in Section 6.3.3.5, the distribution of hormone sensitive metastatic cancer treatments was reflected in the transition probability from metastatic to prostate cancer death, by weighing the treatment of effect of combination therapy according to the relative distribution of treatments. This was also in contrast with the Southampton DAR model,[126] which did not link metastatic treatment distribution to the metastatic treatment effectiveness.

Subsequent metastatic treatment (for hormone-resistant metastatic cancer) was also considered in the York model, for the proportion of individuals who survived the first two years in the metastatic health state (see Section 6.3.4.2). The treatments considered included monotherapy with abiraterone, docetaxel and enzalutamide, and best supportive care, and the treatment distribution was conditional on the type of treatment received at first line (i.e., for hormone-sensitive metastatic cancer). We sourced the hormone-resistant metastatic treatment distribution from the Southampton DAR model[126] (see Table 94, Appendix 11). While the hormone-sensitive metastatic treatment distribution is linked to treatment costs, treatment effectiveness, and adverse events costs, the hormone-resistant metastatic treatment distribution in both models is applied only to estimate the costs of metastatic treatment. While this structural decision was not justified in the Southampton DAR,[126] we considered that extending the model to establish these additional links would be of limited value to this assessment. Therefore, the York model does not consider the effectiveness and safety of hormone-resistant metastatic treatment.

Given that the distribution of metastatic cancer treatments was informed by assumptions, these parameters were not set up probabilistically (i.e., probability distributions were not fitted to these parameters).

6.3.5 Adverse events

6.3.5.1 Biopsy procedure related adverse events

The biopsy procedure is associated with adverse events such as urinary retention, infections, sepsis, haematuria, and death. The cost and HRQoL impacts of these adverse events vary according to their severity and the level of healthcare resource use required to treat them.

The review in Section 4.1 could not establish differences in the type and the rates of adverse events (i.e., the safety profile) between software and cognitive fusion, as well as between different software fusion systems. This was because either comparative safety evidence was not presented, was confounded by the biopsy route of access or the observational nature of the studies limited the ability to attribute differences to the intervention. Furthermore, there is a clear biological mechanism (e.g., clear difference in the number of cores for each MRI-influenced method or a marked increase in procedural time that might increase the likelihood of adverse events from anesthesia) that suggests the safety profile of cognitive and software fusion is different.

The Southampton DAR[126] modelled differences in safety profile between biopsy procedure by route of access and type of anaesthesia. In their revised base-case, the biopsy complications considered for LATP/GATP and LATRUS were mild adverse events (more frequent with transperineal biopsies), adverse events leading to non-elective hospital admission within 28 days of the procedure and peri-procedural death (also within 28 days of the procedure). Transperineal biopsies had a higher rate of

mild adverse events and slightly lower rates of non-elective admission and peri-procedural death.[126, 160] Table 95 in Appendix 11 summarises the adverse event rates and sources used to parameterise the current report base-case analysis (which correspond to the revised base-case estimates in the Southampton DAR).[126]

We note that the adverse event rates estimated for the Southampton DAR[126, 160] did not distinguish between biopsies in terms of sample collection method, so it is unclear whether these estimates are reflective of the safety profile of systematic, targeted or combined biopsies. In the base-case, we assume that the biopsy safety parameterisation of the Southampton DAR is applicable to targeted biopsies and that there are no differences in biopsy complications between these and combined biopsies; this assumption is relaxed in sensitivity analysis.

Parameter uncertainty in the adverse event rates was modelled by fitting beta distributions to these parameters.

6.3.5.2 Localised and locally advanced disease treatment adverse events

Individuals diagnosed as having prostate cancer will receive treatment for localised and locally advanced disease (active surveillance or radical treatment) in the long-term model according to their diagnosed CPG category, while those diagnosed as not having the disease are assumed to receive monitoring (see Section 6.3.4.1). Both radical and conservative (active surveillance) treatment are assumed to have associated adverse events.

In line with the Southampton DAR and the NICE NG131 model,[126, 133] our base-case analyse includes the following categories of adverse events for radical and conservative treatment: i) erectile dysfunction, ii) urinary incontinence, and iii) bowel dysfunction. The rates of adverse events for radical prostatectomy, radiotherapy, and active surveillance were sourced from Table 64, in the Southampton DAR,[126] which was informed by a single trial comparing all three treatments (ProtecT trial[148]). While all patients receiving radiotherapy are assumed to also received ADT (see Section 6.3.4), the Southampton DAR assumed no adverse events from hormone therapy;[126] we also applied this assumption in the York model.

Parameter uncertainty in the adverse event rates was modelled by fitting beta distributions to these parameters.

6.3.5.3 Metastatic disease treatment adverse events

Similarly, to the Southampton DAR model[126] we only modelled adverse events of treatment for hormone-sensitive metastatic disease. Adverse event rates per type of adverse event were sourced from Table 64, in the Southampton DAR,[126], which obtained the rates from three pivotal trials[59, 158, 161] comparing ADT alone to each of the three combination therapies modelled.

Parameter uncertainty in the adverse event rates was modelled by fitting Beta distributions to these parameters.

6.3.6 Health-related Quality of Life

1.1.1.1 Biopsy procedure disutility

The model considers the disutility of biopsy related adverse events. In line with the Southampton DAR[126] a disutility weight was attributed to each type of adverse event (mild, leading to non-elective hospital admissions and death) and then adjusted for duration of the event to generate a QALY loss per type of adverse event. The biopsy procedure QALY loss in the model is then adjusted to reflect the different safety profile between transperineal and transrectal biopsy. The disutility weights and adverse event duration per type of adverse event are reported in Table 96 in the Appendix 11, and where sourced from the Southampton DAR.[126] We did not consider parameter uncertainty in the disutility weights or adverse events duration inputs, given lack of information on their variance.

6.3.6.1 Health state utilities and treatment disutilities

Health state utilities and treatment disutilities were applied as per the Southampton DAR, [126] but adapted for the delayed radical treatment at 2 years in the model for misdiagnosed cases.

6.3.7 Resource use and costs

The resource use and costs considered in the diagnostic pathway include those associated with the biopsy procedure and its adverse eves. The long-term model quantifies the costs of monitoring individuals following the diagnostic procedures in the diagnostic model, the costs of prostate cancer treatment and end of life. Costs in the model are expressed as 2020/21pound sterling and discounted at a 3.5% annual discount rate.

The resource use and cost in the long-term model (costs associated with monitoring, prostate cancer treatment, treatment adverse events and end of life) was largely informed by the Southampton DAR,[126] as were the unit costs sources (updated or inflated to 2020/21 price year as appropriate). Therefore, descriptions of these categories of cost and resource use are brief and refer back to the Southampton DAR model.[126] Emphasis is put into describing elements where our assumptions and/or parameter sources differ from those of the Southampton DAR model.[126]

Uncertainty in resource use and costs inputs was not considered for the large majority of the inputs due to lack of information on their variance and the reliance on assumptions to define parameter quantities.

6.3.7.1 Biopsy procedure costs

This section details the costs associated with the biopsy procedure, which include the following components:

i. Cost of the software fusion system – costs of the fusion software and, in some cases, a workstation (or cart). This cost only applies to the diagnostic strategies which include a software fusion component.
ii. Cost of the ultrasound – cost of the ultrasound probe/transducer, and any required software. This cost applies to diagnostic strategies with either software or cognitive function components, but some software fusion systems are not compatible with third-party ultrasounds.
iii. Cost of software fusion system installation – cost of connecting the software fusion system to the NHS trust IT system; This cost only applies to the diagnostic strategies which include a software fusion component.
iv. Cost of software fusion system maintenance – costs of service contracts to maintain the technology and keep software up to date. This cost only applies to the diagnostic strategies which include a software fusion component.
v. Costs of software fusion system training – staff time costs required to train NHS professionals to perform biopsies. The use of software fusion methods requires additional training compared to cognitive fusion, but the cost of training also varies across biopsy approaches (by route of access).
vi. Cost of staff time to perform the biopsy procedure – cost of urologists, nurses and anaesthetist (for procedures requiring general anaesthesia). This cost varies across biopsy approaches (by route of access and type of anaesthesia), but there is also a difference in procedural time between software fusion and cognitive fusion.
vii. Cost of the biopsy setting – costs of the setting in which the biopsy procedure takes place (outpatient room, theatre session); it varies by route of access, type of anaesthesia, and MRI-influenced method.
viii. Costs of other biopsy devices and consumables – cost of a) devices and equipment (e.g., freehand needle positioning devices, lithotomy beds and biopsy guns), and b) needles and other materials requiring replacement (immediate or after a certain number of uses). These costs are often specific to the biopsy approach (transrectal or transperineal [stabilised, freehand or double freehand]), and may differ across MRI-influenced methods and across software fusion systems, due to compatibility issues.
ix. Cost of histopathology analysis and report– costs of processing the biopsy sample and communicating the results to the patient in a consultation. This cost applies to all strategies but may differ for strategies using different sampling methods (combined vs. targeted-only

biopsy), as these may result in different number of cores being sampled. These costs are reported in the Appendix 11, as they are not software fusion specific.

In the subsequent sections we start by discussing patient throughput and then provide more detail in each component of cost described above, with emphasis in those costs that vary by MRI-influenced method and/or across software fusion systems. Further information is provided in Appendix 11. All costs presented are exclusive of VAT, unless otherwise stated.

Patient throughput

The annual patient throughput represents the average annual number of targeted biopsies (alone or in combination with systematic biopsy) per NHS trust. The annual patient throughput is determinant to calculate the cost of biopsy. The EAG did not identify a source directly reporting this estimate. The evidence considered and the calculations used to inform our base case assumption of throughput are described in Appendix 11. The evidence considered suggests the average annual number of targeted biopsies per NHS trust in England is in a range within 168 and 300. We considered that the expected patient throughput is likely to be closer to the upper bound of the estimated range and consider an annual throughput of 300 targeted biopsies in the base-case analysis (this parameter is varied in sensitivity analysis, see Appendix 12).

Cost of the software fusion system and ultrasound components

The MRI fusion systems under comparison differ in terms of their compatibility with third-party ultrasound devices (and are, therefore, sold without an ultrasound component), with the ultrasound component being an integral part of the software fusion system for some technologies (e.g., KOELIS Trinity). Therefore, the capital costs of the software fusion systems and ultrasound components are reported jointly in this section.

Only five companies provided information on the costs of the technologies under comparison; these were BK Medical UK Ltd (with MIM Software Inc. for bkFusion), Exact Imaging (for FusionVu), Focal Healthcare (for Fusion Bx 2.0), KOELIS (with Kebomed for KOELIS Trinity), and MedCom (BiopSee). No information was provided for the costs of ARTEMIS, iSR'obot™ Mona Lisa, and UroNav Fusion Biopsy System. The capital costs of the software fusion systems and ultrasound components for bkFusion, FusionVu, Fusion Bx 2.0, KOELIS Trinity, and BiopSee, are summarised in Table 43, alongside the lifespan of the equipment.

Table 43 Costs of software fusion system and ultrasound components

Type of software fusion system	Technology	Software fusion costs	Ultrasound costs	Lifespan (years)
Fully integrated system	bkFusion	Cart and software: 52,250*	bk3000 ultrasound: £37,500 Prostate procedural application: £1,800	8 (4 for transducer & sensor clamp)

			DICOM standard with encryption: £1,700 Leakage test kit: £332 Transducer: £15,000 Sensor clamp for the transducer: £200	
	FusionVu	£124,958**		5
	KOELIS Trinity	£23,620	Ultrasound: £45,000, Transrectal software: £39,948+ Transperineal software: £41,754+	5
Compatible with 3rd party ultrasounds	BiopSee	Transrectal: .Software: £15,000 .Cart: £12,000 Transperineal: .Software: £20,000 .Cart for stabilised biopsy: £8,000 .Cart for freehand biopsy: £20,000	NA	10
	Fusion Bx 2.0	Software: £24,244*** Cart: £96,974***	NA	10

*Cost provided for transperineal biopsy only; **Costs originally include value accrued tax at 20%; ***Costs originally expressed in US dollars and subsequently converted to pound sterling at a rate of 0.80812 (represents the average exchange rate between 12/03/2022 to 06/09/2022);[162] +We note that the cost of the transrectal software was reported inconsistently in the company's response to the EAG's additional request for information (in Table and response 7) as £39,431 and £39,948, and the cost of the transperineal software as £42,258 and £41;754. The values used in the model were taken from Table 1 of the company's response to the EAG's additional request for information. DICOM, Digital Imaging and Communications in Medicine; NA, not applicable.

For three software fusion systems (bkFusion, FusionVu, and Koelis Trinity) the software fusion component is integral to the ultrasound component (or the micro-ultrasound component for FusionVu). In the other two systems (Fusion Bx 2.0 and BiopSee) the fusion software is installed on a standalone workstation (or cart), wich is integral to the software fusion system, but does not comprise an ultrasound system. Fusion Bx 2.0 and BiopSee require third-party ultrasounds and transducers to perform prostate biopsies, for which the costs are not reported in Table 43 (as sold by third-party). Both Fusion Bx 2.0 and BiopSee include a cart; the cart is an integral part of each technology.

For software fusion systems that are compatible with third party ultrasounds (i.e., BiopSee and Fusion Bx 2.0), we assume the same cost for the ultrasound components as for cognitive fusion. In the base-case, this cost was derived from the cost of the three standalone ultrasound machines in the Southampton DAR[126] (FUJIFILM transducer and Ultrasound System (inflated to 2020/21 price year according to the NHS cost inflation index [NHSCII];[163] BK ultrasound system and urology software with transducer; Trinity® 3D Prostate Suite plus KOELIS Sidefire Ultrasound probe). We averaged across the costs of these three technologies (with costs updated based on the information provided by bkMedical and KOELIS and Kebomed in the context of the current DAR for the BK ultrasound and Trinity ultrasound components) to estimate an average annual capital cost for ultrasound of £10,846 and £10,974 for transrectal and transperineal biopsies, respectively.

For bkFusion, lifespan estimates provided by the company for the transducer (3-5 years) and leakage test kit (8 years) are said to be end user dependent. We assumed that the transducer lifespan corresponded to the midpoint of the range provided by the company (i.e., 4 years). The lifespan of the sensor clamp for the transducer was not provided by the company despite the EAg request to provide this information for all components, so we assumed it was the same as for the transducer.

Commercial discounts may be available for bkFusion. The company stated that *"We have a 5 years fixed service contract (excluding any civco and mim products) called Priority Care -at point of sale if the service contract is purchased we provide a 10% discount to the priority care quote. If priority care is not purchase within the systems first 2 years of life you can not access this contract again. Alternative contracts are available"*. However, the company did not detail what was included in the Priority Care quote and how much it costs. It is also not clear if the discount applies to maintenance or equipment costs, as we do not know what is covered by the Priority Care quote. Therefore, the information provided by the company is insufficient to implement the discount in the model, and this is not considered by the EAG.

FusionVu uses micro-ultrasound technology, and therefore, does not require ultrasound components. The cost presented for this technology in Table 43 reflects the cost of all equipment and software components. We note that the company stated that they are willing to offer a discount to the UK NHS but that they could not finalise it within the timelines of this DAR.

KOELIS and Kebomed also stated that they can offer discount for multi-unit purchases of KOELIS Trinity, but these depend on the number of units purchased, method of purchase and specification of the units (response to EAG's RFI, question 11). The company did not provide further details on the discounts available, and therefore this is not modelled.

Commercial discounts may also be available for BiopSee according to MedCom, who states that these discounts are usually handled by distributors. As no further information on the applicability and size of the discounts was provided we could not model discounted costs for BiopSee.

The costs of some software fusion systems and/or ultrasound components were specific to the biopsy approach in terms of route of access (transrectal or transperineal) and/or the fixation method (stabilised, [single] freehand and double-freehand), so for KOELIS Trinity and BiopSee costs will vary conditional on the diagnostic strategy they are being used in. The costs provided by the company for bkFusion were reported solely for a transperineal procedure, despite the EAG request to provide costs by biopsy approach. Therefore, it was assumed that the costs of bkFusion are the same across biopsy approaches.

The software costs assumed for Fusion Bx 2.0 (£24,244) assume the purchase of a perpetual license. The company also provided the cost of an annual license costing a third of the perpetual license. Given the lifespan of Fusion Bx 2.0 exceeds three years (point beyond which the annual cost of a perpetual license becomes lower than the annual license), we did not consider annual licenses as an option. The company stated that a discount on the software and hardware components of Fusion Bx 2.0 of up to 30% could be offered to the NHS, depending on the number of systems purchased. We did not implement this discount on our base-case analysis, as the company did not specify the level of discount applied conditional on number of units purchased.

Some software fusion systems had optional probe holders and software components, which were not considered in the costs of the ultrasound components, as these are not essential components of the technology. We note that the cost of Fusion Bx 2.0 includes one probe holder as an integral part of the system, and therefore, this cost was not excluded.

The costs of software fusion systems and ultrasound components were annuitized at a 3.5% discount. Annuitized costs and costs per biopsy are reported in Table 44

Table 44 Costs of software fusion and ultrasounds components

		Annuatized cost			**Cost per biopsy**		
	Biopsy approach	**TR**	**TP stabilised**	**TP freehand**	**TR**	**TP stabilised**	**TP freehand**
Technology	**Type of system**						
bkFusion	Fully-integrated*	£17,152			£57.17		
FusionVu	Fully-integrated*	£26,740			£89.13		
KOELIS Trinity	Fully-integrated*	£23,233	£23,619	£ 23,619	£77.44	£78.73	£78.73
BiopSee	Software fusion alone	£13,982	£14,227	£15,621	£46.61	£47.42	£52.07
Fusion Bx 2.0	Software fusion alone	£24,928	£25,057		£83.09	£83.52	
bkFusion	Cognitive fusion	£10,846	£10,974		£36.15	£36.58	

*Includes the cost of each technology own brand ultrasounds components. TP, transperineal; TR, transrectal

Cost of installation of software fusion systems

One company (Medcom) reported the time required to install the MRI fusion system technology to the NHS trust IT system as ranging between 30 to 60 minutes. We assumed that this results in a one-off

staff time cost, which is applicable to all software fusion systems. The cost of installation was estimated assuming it would take 45 minutes (midpoint of the time range provided by Medcom) of an IT worker time (costed at £35.67 per hour [average working hour of band 4 hospital-based scientific and professional staff][163]). The cost was distributed over the annuitized (3.5% annual rate) average lifespan of the five software fusion systems for which the companies had submitted costing information. The resulting annual cost and cost per biopsy were estimated to be £3.97 and £0.01, respectively.

Cost of software fusion system maintenance

The costs of maintaining the software fusion systems mostly consist of the costs of service contracts. These contracts also include maintenance of the ultrasound components when the ultrasound components are integral to the software fusion system. The maintenance contracts are summarised in Table 45, alongside the annual cost estimate applied in the model.

Table 45 MRI fusion system maintenance contracts

Technology	Maintenance contract duration and cost	Costs of maintenance applied in the model	
		Annual cost	**Cost per biopsy**
bkFusion	5 years: £66,975.00	£13,395.00	£44.65
FusionVu	NR	£12,206.12	£37.20
KOELIS Trinity	Essential - 1 year: £5,500.00	£11,017.24	£29.76
	Comfort - 1 year: £7,465.52		
	Serenity - 1 year: £11,017.24		
BiopSee	NR	-	-
Fusion Bx 2.0	1 year: £9,697.44*	£9,697.44	£32.32

*Costs originally expressed in US dollars and subsequently converted to pound sterling at a rate of 0.80812 (represents the average exchange rate between 12/03/2022 to 06/09/2022);[162] NR, not reported.

Three companies provided information on the duration and cost of the maintenance contracts. Most contracts had an annual duration; only bkFusion had a five years maintenance contract. Given the lifespan of bkFusion is greater than 5 years, we distribute this cost equally over time and apply it as a cost of £13,395 per annum in the model. We note that there are discounts available for alternative maintenance contracts for bkFusion, which have eben described above.

KOELIS Trinity has three levels of maintenance contract, which differ in terms of annual costs; the levels are: Essential, Comfort and Serenity. According to the company Serenity is the level most often purchased (50%) followed by Essential (34%), and Comfort (16%). In the model, we assume the annual cost of the contract to be a weighted average of the three contract levels by the corresponding "market share", resulting in an annual cost and a cost per biopsy of £8,926.90 and £29.76, respectively.

We note that the maintenance service contract cost for Fusion Bx 2.0 is an approximate estimate provided by the company, who stated that this would typically cost $12,000US or less and that they plan to enlist a UK distributor to perform this service. Alternatively, the maintenance could be conducted by hospital staff who are responsible for performing preventative maintenance of other medical devices, and who would need to undergo annual on-site maintenance training (1-2 hours). The company also stated in the responses submitted to NICE that if the maintenance contract is longer than one year, the cost would be discounted accordingly. However, this does not provide information on the level of discount over time, so this potential discount cannot be implemented without more detail. In the model, we consider only the approximate cost of an annual maintenance service delivered by the company or their distributors.

No maintenance costs were provided for FusionVu and BiopSee. The company who commercialises FusionVu, stated that their technology is serviced through a local distributor in the UK under annual or more contracts, but could not yet provide a cost estimate for the contract. Therefore, we have assumed that the FusionVu maintenance contract costs is an average of the two software fusion systems with fully integrated software fusion system and ultrasound components (bkFusion and KOELIS Trinity). Medcom stated that BiopSee does not require any maintenance, as damaged parts can be repaired on demand and reported cost range for repairing accessory equipment (e.g., £200 - £600 to replace a mouse or an accidentally damaged cable, and £100-£3,000 to replace a damaged stepper or stabiliser). However, it is unknown how often damage to different components is likely to occur, and so estimate a maintenance cost on a per damaged part basis. We could have assumed a maintenance cost similar to that of Fusion Bx 2.0 (the other software fusion system that does not have an integral ultrasound component), however we note that the maintenance cost for Fusion Bx 2.0 is an approximate estimate provided by the company and assumes that there is a service contract (not available for BiopSee). In our base-case analysis, we assume that there is no cost attached to maintaining BiopSee.

The cost of software updates is included within the maintenance contract for most technologies. One of the exceptions is bkFusion, which only includes software malfunction fixes. No software update costs were provided for bkFusion, but we note that the lifespan of the hardware and software components for this technology are generally the same. The cost of software updates for BiopSee software fusion software, is 50% of the software cost and new versions are usually released annually, according to the company. The company also stated they do not plan to withdraw the current version being used in the UK NHS. Therefore, we assume that no additional costs for software updates need to be considered for any of the technologies.

Cost of MRI fusion system training

The technology specific cost of software fusion systems training consists of the cost of staff time to attend the training sessions, as companies do not charge for training provision. Each company provides a core training programme composed of different elements. The information provided by the companies on the NHS staff who should undergo core training and the time required per training component is summarised in Table 98, in the Appendix 11.

To estimate an annual cost of training for the use of software fusion per trust for each technology, we assumed that core training would be delivered to two urologists, two nurses, one radiologist and one sonographer, and training would remain up to date for 5 years (the shorter lifespan across fusion software). We used the training duration provided by the companies to estimate staff time requirements, and assumed the same time for all categories of staff unless the company stated different times by category of staff. Where the companies provided training duration as a range we assumed the staff time requirement would correspond to the midpoint of that range. We did not include any staff time for theatre list or support to clinical cases, as we assumed that this would not result in additional time requirements in relation to the procedure time. The information used to estimate the costs of software fusion systems training is presented in Table 46, alongside the annuitized annual cost (at 3.5% per annum) of training for each technology. Unit costs were sourced from the PSSRU (2021) unit costs report.[163]

We also considered the cost of training to perform biopsy procedures more generally; these were assumed to vary by biopsy access route (transperineal vs. transrectal) in line with the Southampton DAR. We assumed the same level resource use per biopsy approach as was assumed in the Southampton DAR and updated the unit costs to reflect our analysis price year.[163]

Table 46 MRI fusion systems training costs

	NHS staff time				Annuitized annual cost	Cost per biopsy	Unit cost	
	Urologist	Nurse	Radiologist	Sonographer			Staff	Cost per working hour
Technology specific								
bkFusion	2 x 11.25 hours*	2 x 11.25 hours*	1 x 11.25 hours*	1 x 11.25 hours*	£1,029.97	£3.43	Urologist & radiologist	£87.50, average of hospital-based registrar and medical consultant PSSRU (2021)[163]
FusionVu	2 x 3 hours	2 x 3 hours	1 x 3 hours	1 x 3 hours	£274.66	£0.92		
KOELIS Trinity	2 x 3 hours	2 x 3 hours	1 x 3 hours	1 x 4 hours**	£285.86	£0.95	Nurse	£46.00, average of hospital-based nurse specialist/team leader (band 6) and nurse advanced/team manager (band 7) PSSRU (2021)[163]
BiopSee	2 x 3 hours	2 x 1 hour	1 x 3 hours	1 x 1 hour	£203.90	£0.68	Sonographer	£52.33, average hospital-based scientific and professional staff (band 6) PSSRU (2021)[163]
Fusion Bx 2.0	2 x 1.625 hours	2 x 1.625 hours	1 x 1 hour***	1 x 1.625 hours	£137.07	£0.46		
Biopsy approach specific								
Transrectal	5 x 1 hour	-	-	-	£437.50+	£1.46	Urologist	£123, medical consultant PSSRU (2021)[163]
Transperineal	5 x 8 hours	-	-	-	£3,500.00+	£11.67		

*Assumes a working day corresponds to 7.5 hours; **Assumes installation training is only undertaken by one sonographer; ***Assumes that phantom prostate biopsy training is not undertaken by the radiologist; +, not annuitized as this was estimated as annual training requirement in the original source; PSSRU, Personal Social Services Research Unit.

FusionVu has a free-of-charge optional training programme, the Mastery programme. The company did not provide clear information on the staff time requirements to undergo this optional training, or clarify to whom it would be delivered. The company stated that the effectiveness of the Mastery programme was studied in Cash et al., 2022[164], but it was not possible to ascertain based on the information provided if the Mastery programme described by the company corresponded to the training programme assessed in this publication. Given this and the optional nature of this training component, we have not included this cost in our cost-effectiveness analysis.

Cost of staff time to perform the biopsy procedure

The staff costs associated with the biopsy procedure are likely to vary depending on the biopsy route of access and the type of anaesthesia. We based our estimates of staff time requirements to conduct the biopsy on the Southampton DAR.[126]

In the Southampton DAR, each local anaesthesia biopsy was assumed to require one urologist and two nurses, with general anaesthesia biopsy further requiring one anaesthetist. The time required for local anaesthesia transperineal (LATP) biopsy was sourced from the published literature for two devices (0.41 and 0.33 hours for CamPROBE and PrecisionPoint™, respectively), with an average of the two assumed for the devices for which there was no published evidence. For general anaesthesia transperineal (GATP) biopsy, the procedure time was assumed to be one hour. To estimate the procedure time of local anaesthesia transrectal (LATRUS) biopsy, the EAG applied a ratio of procedure time between the transrectal and transperineal approach (0.84) derived from the literature to the time estimates by transperineal device.

In this report, we have assumed the procedure time of LATP conducted with PrecisionPoint™, as the diagnostics consultation document (DCD) for the Southampton DAR[29] suggests that CamPROBE will not be recommended for use in the NHS UK. We applied the same LATRUS/LATP time ratio as in the Southampton DAR to the LATP time estimate to calculate the LATRUS procedure time and assumed GATP would take one hour. We also assumed that 50% of procedures would be undertaken by urologists and 50% by sonographers. The remaining staff requirements were assumed the same as in the Southampton DAR (i.e., two nurses plus one anaesthetist if GATP). We applied the same unit cost as those used to cost training costs (see Table 46) and assumed the same unit cost for anaesthetist time as for the urologist time. The procedure time costs by route of access and type of anaesthesia are summarised in Table 47.

Table 47 Biopsy procedure staff time costs

Biopsy approach	Procedure time (hours)	NHS staff				Cost per biopsy
		Urologist	Nurse	Sonographer	Anaesthetist	
LATP	0.33	0.5	2	0.5	-	£60.36
GATP	1				1	£270.42
LATRUS	0.28				-	£50.70

GATP, general anaesthesia transperineal; LATRUS, local anaesthesia transrectal; LATP local anaesthesia transperineal.

The MRI-influenced biopsy procedure time may further increase when this is performed with fusion software compared to cognitive fusion. This additional time is due to the need to i) contour the prostate in the MRI and ultrasound images, and ii) to connect the MRI fusion system. The companies provided different estimates of how much time would be added to the biopsy procedure when using their technologies (see Appendix 11, Table 99), but these were not supported by published evidence. As discussed in Section 4.5.2, the procedure time estimates in the diagnostic literature do not allow disentangling the additional procedure time due to software compared to cognitive fusion from procedure time differences associated with the biopsy approaches. The clinical advisers to the EAG commented that the additional procedure time for software fusion (vs. cognitive fusion), should be approximately 10 minutes in a high throughput centre, 5 minutes of which would correspond to additional time to import and obtain the appropriate MRI sequences (radiologist time) and 5 minutes during the actual biopsy (urologist/sonographer and nurse time) to connect the software fusion systems and contouring the ultrasound. They also noted that these time estimates could be longer when the use of these interventions is first rolled out, due to lack of experience.

We calculated the additional staff time costs required to conduct software fusion based on the time estimates provided by the clinical advisers to the EAG. We assumed the same staff requirements per type of biopsy approach, as for the core biopsy procedure time, and further accounted for the additional time requirements for one radiologist. We applied the same unit cost as those used to cost training costs (see Table 46) and assumed the same unit cost for anaesthetist time as for the urologist time. The additional procedure staff time costs for software fusion compared to cognitive by route of access and type of anaesthesia are summarised in Table 48.

Table 48 Additional biopsy procedure staff time costs for software fusion

Biopsy approach	NHS staff time (minutes)					Additional cost per software fusion biopsy
	Radiologist	Urologist	Nurse	Sonographer	Anaesthetist	
LATP	1 x 5	0.5 x 5	2 x 5	0.5 x 5	-	£22.53
GATP					5 x 1	£29.83

LATRUS					-	£22.53

GATP, general anaesthesia transperineal; LATRUS, local anaesthesia transrectal; LATP local anaesthesia transperineal.

Cost of the biopsy setting

The Southampton DAR[126] examining the cost-effectiveness of LATP, GATP and LATRUS considered a cost for the setting on which the biopsy took place, with LATRUS and LATP being conducted in an outpatient room, and GATP in a theatre session. These costs were sourced from an unpublished study submitted by the sponsor of one of the technologies under assessment in the Southampton DAR[126] and suggested a unit cost of £43 and £129 per hour for the outpatient room and theatre session, respectively. These unit costs were inflated to 2020/21 price year using the NHSCII,[163] applied to the duration of the procedures for each biopsy approach, to estimate the cost of the setting.

The micro-costing study is not described in sufficient detail to understand what is included in the costs of biopsy setting. This DAR's EAG has decided to include the cost of setting for consistency with the Southampton DAR,[126] but notes the opacity of the cost estimates as a potential limitation.

The cost of biopsy setting applied in the model for strategies using cognitive fusion were calculated by multiplying the time of the procedure by biopsy approach (see Table 47) by the unit costs by setting (inflated to 2020/21 price year)[163] in the Southampton DAR.[126] For strategies using software fusion, we assumed that the procedure would take 10 additional minutes (in line with the assumptions to estimate the additional staff time to conduct software fusion and assuming that the MRI is also done in an outpatient setting). Costs associated with biopsy setting by biopsy approach and MRI-influenced method, as well as the model inputs to estimate these are summarised in Table 49.

Table 49 Costs of biopsy setting

Biopsy approach	Procedure time (hours)		Unit cost (per hour)	Cost per biopsy	
	Cognitive fusion	Software fusion		Cognitive fusion	Software fusion
LATP	0.33	0.50	£44.32	£14.63	£22.01
GATP	1.00	1.17	£132.97	£132.97	£155.14
LATRUS	0.28	0.44	£44.32	£12.29	£19.67

GATP, general anaesthesia transperineal; LATP, local anaesthesia transperineal; LATRUS, local anaesthesia transrectal

Costs of other biopsy devices and consumables

The biopsy procedure requires other devices and consumables which may vary by biopsy approaches (GATP, LATP, LATRUS). While these devices and materials are not are required conduct biopsy procedures with either software or cognitive fusion some technologies have compatibility issues

which mean that costs of technology specific materials may have to be considered where appropriate to fully account for differences in costs between the different software fusion systems. For example, FusionVu is only compatible with needle guides commercialised by Exact Imaging, meaning that in principle, you cannot use other needle guides that would be suitable for cognitive fusion or other software fusion systems without compatibility issues. In our base-case analysis, we apply a simplifying assumption that the costs of the biopsy devices do not vary by MRI-influenced methods, but we consider potential differences in scenario analysis

Transperineal biopsies can be conducted with a i) grid and stepper unit; ii) freehand device (the Southampton DAR[126] assessed five of these devices) or iii) coaxial needle (one such device assessed in the Southampton DAR). Grid and stepper units are used for stabilised biopsies with the stepping unit usually fixed to a stabiliser (mounted onto a table or supported by a floor stand). The stepper is a reusable device used to hold the ultrasound probe, while a (single use or reusable) grid is used to guide the needle insertion. Grid and stepper units can be used to perform transperineal biopsies under or local general anaesthesia. Recent LATP techniques are performed using an access needle guide (or equivalent) to pierce the perineum and through which the biopsy needle passes to sample the prostate. These techniques can be performed using i) freehand devices attached to the ultrasound probe or ii) a co-axial needle not attached to the probe (also known as double freehand technique).

We based the costs of a grid and stepper unit for stabilised GATP and LATP biopsy on the estimates used in the Southampton DAR for this cost element with adjustments to reflect our throughput estimates. We assumed a cost of reprocessing reusable materials of £5.15 (cost of cleaning and sterilising), sourced from the Southampton DAR and inflated to 2020/21 price year using the NHSCII.[163] Costs of devices to conduct transperineal stabilised biopsies by software fusion technology and for cognitive fusion are summarised in Table 50, which also summarises the corresponding costs for freehand transperineal and transrectal biopsy.

For costing transperineal biopsies with a (single) freehand techniques with cognitive fusion, we have considered the costs of the five freehand devices evaluated in the Southampton DAR[126] (PrecisionPoint [BXTAccelyon], EZU-PA3U [Hitachi], UA1232 [BK Medical], Trinity® Perine [KOELIS and Kebomed], and SureFire [Delta Surgical]; inflated to 2020/21 price year using the NHSCII[163]). We have updated the costs of Trinity® Perine device based on the cost of the reusable Perine Grid 18G provided by KOELIS and Kebomed in the context of the current DAR (£779.31; 100 uses). We note that KOELIS and Kebomed also commercialise single use Trinity® Perine grids (costed as £62.04 and £86.20 for a Mini grid and a Full grid, respectively – not modelled); this are not included in the model but would yield higher costs per biopsy than the single use devices. We included a £5.15 cost of reprocessing for the reusable devices (EZU-PA3U [Hitachi], UA1232 [BK

Medical], Trinity® Perine [KOELIS and Kebomed]. In the base-case, the cost of freehand devices is an average of the costs for the five devices and applies equally to cognitive and software fusion.

The costs of transperineal devices applied in the model for the base-case analysis are summarised in Table 50.

Table 50 Costs of transperineal biopsy devices

Technology	Cost per stabilised biopsy	Cost per freehand biopsy
Software fusion (base-case)	£90.44	£81.86
Cognitive fusion (base-case/scenarios)		

We did not consider the costs of LATP with double freehand technique, as the provisional DCD for the previous DAR does not recommend the use of double-freehand devices to conduct LATP. We also did not consider any device costs to conduct LATRUS in line with the Southampton DAR.[126].

We included the annuitized cost of a lithotomy bed (£10,308, 10 years lifespan) in the calculations of the cost per biopsy of transperineal biopsy; this cost was sourced from the Southampton DAR,[160] and inflated to 2020/21 price year using the NHSCII.[163]

The costs of general consumables by biopsy approach were also sourced from the previous DAR,[126, 160] where they are detailed (Table 113 of the Southampton DAR). We applied a cost per biopsy of £80.7, £65.55, and £79.10 for LATP, GATP and LATRUS, respectively.

Cost of histopathology analysis and report

The Southampton DAR[160] assumed that the cost of histopathology analysis was dependent on the number of cores sampled and each biopsy involved sampling 12 cores.

There was limited comparative evidence to inform any differences in number of cores sampled between cognitive and software fusion identified in the clinical review, as most diagnostic accuracy studies performed a fixed pre-specified number of cores per biopsy. One RCT[31] reported the median number of cores per subject undergoing a targeted biopsy; 4 (IQR:3-5, n=79) and 3 (IQR: 3-3; n=78) for software and cognitive fusion, respectively. This suggests that fewer cores than 12 cores would require analysis per targeted biopsy, and that differences between MRI-influenced methods are small. However, the study had a small sample size and this was not a primary outcome, so it is unlikely that the study was powered to identify any differences in this particular outcome between MRI-influenced methods.

The unit cost of histopathology analysis of the cores sampled through biopsy was sourced in the Southampton DAR[126, 160] initially from a histopathology pricing document by the University of Surrey, and then corrected to an HRG cost (£36.58; currency code DAPS02: Directly Accessed Pathology Services - Histopathology and histology).[165] The resulting cost for the analysis of a 12-core biopsy was £438.96 in the Southampton revised base-case analysis, which assumed the unit costs applied to each core tested. This level of resource used applied is more in line with some systematic biopsies (Section 2.2.3).

In the York model, we assumed that the NHS reference cost applied in the Southampton model, applied to a single targeted biopsy (with fewer than 12 cores sampled per biopsy). We sourced the same HRG currency cost (£16.29) from the latest version of the NHS reference costs[166] and applied it to each targeted biopsy. We also did not identify comparative evidence on the number of cores sampled for targeted and combined biopsies, so no differences were assumed. We note that if we have underestimated the histopathology analysis cost of biopsy (targeted or combined), this would only be likely to impact the cost-effectiveness estimates if there were considerable differences in the rates of subsequent biopsy between the intervention and comparator.

We also considered the cost of reporting to the patient the biopsy result. In line with the previous DAR,[160] this was assumed to require a 30 minutes appointment with an urologist (medical consultant, £123 per hour)[163], resulting in a cost per biopsy of £61.50.

Costs per software and cognitive fusion biopsy

Table 51 summarises the aggregated cost per biopsy for each technology and by biopsy approach, with further breakdown of costs in Table 101, Table 102, and Table 103 in the Appendix 11.

Table 51 Cost per biopsy by technology and biopsy approach

		Biopsy approach		
	Technology	**LATRUS**	**GATP**	**LATP**
Technology specific	bkFusion	£147.48	£380.67	£231.68
	FusionVu	£169.47	£402.67	£253.68
	KOELIS Trinity	£150.37	£384.86	£235.87
	BiopSee	£89.51	£323.52	£179.18
	Fusion Bx 2.0	£158.10	£391.72	£242.73
	Cognitive fusion	£48.44	£260.00	£133.07
Non-technology specific		£209.95	£634.15	£239.25
Total cost per biopsy	bkFusion	£356.53	£914.82	£470.93
	FusionVu	£378.53	£936.82	£492.93
	KOELIS Trinity	£359.43	£919.01	£475.12
	BiopSee	£298.56	£857.67	£418.43
	Fusion Bx 2.0	£367.15	£925.87	£481.98
	Cognitive fusion	£257.49	£794.15	£372.32

GATP, general anaesthesia tranperineal biopsy; LATP, local anaesthesia transperineal biopsy; LATRUS, local anaesthesia transrectal biopsy

As stated in Section 6.3.2, we assumed for the first biopsy in the diagnostic pathway 65% of biopsies were conducted with LATP and the remainder with LATRUS. For the repeat biopsy, we assume that 60% are LATP, 30% LATRUS and 10% are GATP (to reflect those individuals where there was concern that first biopsy may not have been accurate due to patient moving excessively during the procedure). We weighted the costs per biopsy approach by the corresponding proportions for first and repeat biopsy to estimate their costs in the model; these costs are reported in Table 52.

Table 52 Cost of first and repeat biopsy in the model

Technology	1st biopsy cost	Repeat biopsy cost
bkFusion	£430.89	£481.00
FusionVu	£452.89	£503.00
KOELIS Trinity	£434.62	£484.80
BiopSee	£376.47	£426.39
Fusion Bx 2.0	£441.79	£491.92
Average cost software fusion	£427.33	£477.42
Cognitive fusion	£332.13	£380.05

There are a number of uncertainties in the biopsy procedure costs. These pertain to:

- The set of essential components that are integral part of each technology and the lifespan for all components;
- The potential commercial discounts that may be offered by the companies, what is included in the commercial arrangements and how do these apply to each technology;
- What additional costs may stem from compatibility issues with existing equipments and accessories in use in the NHS;
- The additional time required to perform software fusion;
- How training for the use of software fusion is delivered (to whom and for how long), and if the training requirements differ substantially between software fusion technologies.

Given these uncertainties and that it was not possible to calculate diagnostic performance evidence by individual software fusion devices at the granularity of classification (ISUP G1, ISUP G2, ISUP G3, and ISUP G 4 or 5) required by the economic model, it was considered the biopsy procedure costs for each individual technology was potentially misleading to decision makers. Thus, we apply the average biopsy cost across all software fusion tehcnologies in this assessment for which cost data was submitted by the companies. Given this and the numerous uncertainties in the cost estimation of each software fusion technology, it was not considered appropriate to compare each software fusion

technology against each other and cognitive fusion in the model. Instead, in the base-case analysis, we apply the average cost per biopsy across all software fusion. Individual software fusion technology costs are presented alongside the base-case analysis results to illustrate how their individual costs would impact on the estimates of cost-effectiveness.

6.3.8 Biopsy procedure adverse events costs

The biopsy procedure related adverse event costs considered in the diagnostic pathway model were estimated by multiplying the adverse event rate by the unit cost for each type of adverse event. The unit costs for each type of adverse event were derived from the Southampton DAR sources (updated for the 2020/21 price year by either using the corresponding versions of national tariffs (e.g., PSSRU and NHS reference costs) or inflating costs using the NHSCII,163 as appropriate) and using the same assumptions (e.g., on resource use required to treat a mild adverse events);126 further details are presented on in Table 104 in Appendix 11

6.3.9 Monitoring costs

Routine monitoring costs at model entry apply to all patients who enter the long-term model. In the model the set of monitoring tests and schedule varying according to whether the individuals have been diagnosed:

- Localised and locally advanced prostate cancer, and if so, monitoring also varies with:
    - the diagnosed CPG category (CPG 1, CPG 2-3 or CPG 4-5);
    - treatment assigned (active surveillance or radical treatment;
    - and time in the model (first, second or subsequent years).
- Or not, and if so, monitoring only varies with the underlying true disease status (no prostate cancer or CPG1-5).

Table 105 in Appendix 11 summarises the resource use and cost per year of the monitoring tests considered in the model for patients in the diagnosed as localised (and locally advanced) disease health states. These costs are applied from model entry (cycle 0) and while individuals remain in the localised disease health states.

We assumed that individuals without a prostate cancer diagnosis would also undergo routine monitoring, regardless of whether they had prostate cancer. In contrast the Southampton DAR,[126] only attributed a cost of monitoring to those with localised prostate cancer who had not been identified as having prostate cancer. We changed this assumption in the York model, because in principle these two groups of individuals would be indistinguishable, as true disease status would be unknown to clinicians. We assume that in both groups individuals receive the same monitoring schedule when they are discharged to primary care: an annual PSA test (velocity test at a threshold of 75ng/ml/year)

for up to ten years, performed at a 10 minutes nurse-led appointment, and followed by a cognitive fusion biopsy (costed at £477.75; assumes 35% LATRUS and 65% LATP) if the PSA test results is positive. As per the Southampton DAR,[126] the probability of testing positive in the PSA test for those with prostate cancer was assumed to be 0.69, which corresponds to the sensitivity of the corresponding PSA velocity test used in NICE NG131 model. We further assumed that the probability of testing positive in the PSA test for those without prostate cancer corresponded to one minus the specificity of the same test (1- 0.56=0.44). The testing schedule is similar to what was modelled in the Southampton DAR[126] for those with prostate cancer who were diagnosed as not having the disease, but the first PSA test is assumed in the York model to occur within one year in the long-term model (rather than 6 months). The annual cost per year of monitoring applied in the York model was £342.50 and £223.06 for those with and without prostate cancer respectively. These costs are applied from model entry (cycle 0) and for up to 10 years in the entry health states.

After two years in the model, individuals in the local disease health states are assumed to to be correctly identified at their true disease status, and move to the monitoring regime that matches their true disease status.

Individuals who enter the metastatic health state incur a one-off monitoring cost of 577.83, corresponding to the same resource use as in the Southampton DAR[126] (i.e., one CT and bone scan).

6.3.10 Prostate cancer treatment costs – localised and locally advanced disease

Individuals identified as having localised or locally advanced prostate cancer are assumed to receive treatment at long-term model entrance according to their diagnosed CPG (see distribution of treatments by diagnosed CPG in Section 6.3.4). Individuals who receive active surveillance is assumed to not incur any treatment costs (only monitoring costs as detailed in Section 6.3.9), so costs of treatment are only incurred by those who undergo radical treatment.

Radical treatment resource use and costs vary according to the type of radical treatment (radical prostatectomy, external radiotherapy or brachytherapy). The cost of each type of radical treatment procedure applied in the York model are reported in Table 106 Appendix 11, alongside details on resource use and unit costs. We note that the cost of brachytherapy has increased considerably in relation to the one used in the Southampton model (£9,156.96 vs. £3,106.02); these differences are driven by an increase in the unit cost of delivering brachytherapy in an outpatient setting (as well as increased activity for the corresponding currency code) in 2020/21 compared to 2019/20. The costs of the radical treatment procedures were applied as one-off costs at long-term model entry (cycle 0). For those who were misdiagnosed and treated with conservative treatment, it is assumed that they receive radical treatments according to their true disease status after 2 years in the model.

In addition to the medical procedures, we also included the cost of ADT for those patients who treated with radiotherapy, according to NICE guidance.[10] ADT in the localised disease setting was assumed to consist of the same treatments as in the Southampton model, i.e., bicalutamide 50mg for 21 days followed by luteinising hormone-releasing hormone agonists (LHRH) (either leuprorelin 11.25 mg [every 3 months], triptorelin 11.25mg [every 3 months] or goserelin 3.6mg [every 28 days]). Similarly, to the Southampton model, the duration of LHRH treatments was varied according to category of prognostic risk; we assumed LHRH treatment duration would be three and six months for those diagnosed in CPG1 and CPG 2-3 categories. For those diagnosed in the CPG 4-5 category, we updated the duration of treatment for 3 years, in line with the current NICE guidance for that prognostic risk group (see Section 2.2).[10] The costs of ADT included drug acquisition and administration costs and were costed as per the Southampton DAR (updated to the current price year).

6.3.11 Prostate cancer treatment costs – metastatic disease

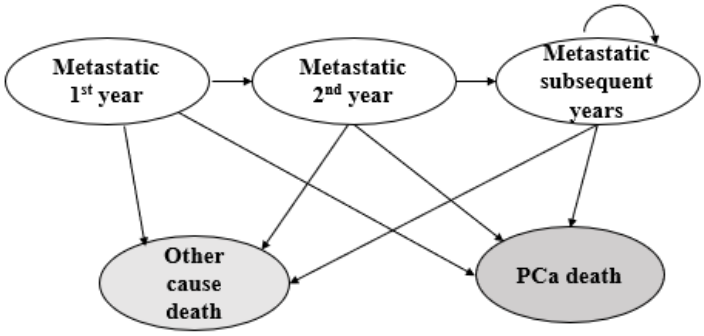
Metastatic disease treatment was assumed to consist of hormone-sensitive disease treatment for the first two years in the metastatic health states, followed by hormone-resistant disease treatment. Cost of metastatic treatment are summarised in Table 107, Appendix 11; these include drug acquisition and administration costs.

In line with the Southampton DAR, hormone-sensitive metastatic treatment was modelled as a blended treatment consisting of ADT alone (but not identical to the regimes described for the localised disease setting, as course of bicalutamide 50 mg is longer), or in combination with either docetaxel, apalutamide or enzalutamide. We updated the distribution of treatments for the hormone-sensitive metastatic treatment, as described in Section 6.3.4.2. We note that yearly costs of metastatic treatment have increased considerably in the York model compared to the Southampton model (e.g., metastatic hormone-sensitive first year cost increase to £15,603.87 from £8,388.63), as a consequence of the increased proportion of individuals treated with ADT combined with enzalutamide, due to the high cost of enzalutamide. Furthermore, although we apply the same docetaxel treatment regimen (i.e., six cycles [delivered every three weeks] at a dose of 75 mg/m^2; body surface area 1.91) as in the Southampton model, in the York model the two-year docetaxel treatment costs are assumed to be distributed evenly between two model cycles (constant annual cost).

The Southampton DAR states that ADT alone or in combination was taken until disease progression, which was assumed to occur after two years. We also make their stated assumption but we implemented it in a different way. In the Southampton model, a cost for first and second year is estimated for metastatic treatment (both treatment for hormone-sensitive and hormone resistant disease) and applied to individuals in the metastatic disease state at first and second year (modelled in a way akin to tunnel states), respectively. Thus, in the Southampton model the cost of hormone-

sensitive and hormone resistant treatment is applied to the same set of individuals. In the York model, we explicitly model a set of three tunnel health states representing the first, second and subsequent years of metastatic disease (Figure 11). We applied the costs of hormone-sensitive metastatic treatment to individuals in the first and second year of metastatic tunnel health states, and the costs of hormone-resistant metastatic treatment costs are applied as a one-off cost to individuals who enter the 'metastatic subsequent years' health state.

Figure 11 Diagram of metastatic tunnel health states and death states



Another difference between models, is that in the York model metastatic treatment (and monitoring) is assumed to apply to all patients with metastatic disease, as we do not distinguish between diagnosed and undiagnosed metastatic disease (the latter does not appear to incur treatment costs in the Southampton model). Thus, we implicitly assume that all individuals with metastatic disease have been diagnosed.

All metastatic treatments costs are applied as an average of the costs of the different type of treatments weighted by their treatment distribution (see treatment distribution in Table 42, Section 6.3.4).

6.3.12 Prostate cancer treatment adverse event costs – localised and locally advanced disease advanced prostate cancer

The model considers the costs of managing the adverse events from active surveillance, radical prostatectomy and radiotherapy for localised or locally advanced prostate cancer. These costs were estimated by multiplying the adverse event rates (see 6.3.5.2) by the unit cost of the corresponding adverse event (Table 108, Appendix 11). The costs are applied in the model as a one-off at localised and locally advanced health states to the proportion of patients who receives each treatment (see treatment distribution in Table 42, Section 6.3.4).

6.3.13 Prostate cancer treatment adverse event costs – metastatic prostate cancer

The costs of managing metastatic treatment related adverse events was applied in the model. Similarly, to the Southampton model only adverse events of treatment for hormone-sensitive

metastatic disease were included. The adverse event costs for androgen therapy alone and in three alternative combinations (with docetaxel, apalutamide or enzalutamide) were estimated by applying the unit cost per type of adverse event (Table 108, Appendix 11) by the corresponding rate (see Section 6.3.5.3). The resulting costs per treatment were then applied as a one-off cost at entrance to the 'metastatic 1st year' health state. The one-off cost was estimated by weighing each treatment cost by the metastatic treatment distribution (see treatment distribution in Table 94 in Appendix 11).

6.3.14 End of Life Costs

End of life costs are applied to all individuals who die in the model of other cause or prostate cancer death, but not to those who have died of peri-procedural biopsy complications (see Section 6.3.5). This one-off cost is applied to individuals who enter the death states at each cycle in the model, and it was sourced from Round (2015)[149] and inflated to 2020/21 price year.[163]

6.4 Analytic Methods

6.4.1 Overview

The diagnostic and long-term model is evaluated deterministically and probablistically for the base-case analysis (1,000 Monte Carlo simulations) so as to incorporate the joint parameter uncertainty across all of the model inputs according to the probability distributions assigned to each. The parameters set up probabilistically in the model are identified in Table 109 in Appendix 11.

Following conventional decision rules for cost-effectiveness, the mean costs and QALYs for the two strategies (cognitive or software fusion) for two set of comparisons (targted biopsy alone or combined with systematic biopsy) are presented and cost-effectiveness compared by estimating the ICERs, as appropriate. A net health benefit (NHB) approach is also applied, for which the unambiguous decision rule. Net-benefits can be expressed on the effect scale (Net Health Benefits; NHB), which is calculated at the two cost-effectiveness thresholds at the lower and upper bound of the range used by NICE to guide decision making (i.e., £20,000 and £30,000 per additional QALY). The formula to estimate net health benefits is presented below:

$$Net\ Health\ Benefit\ (NHB) = QALYs - \frac{Costs}{Cost{-}effectiveness\ threshold}$$

Heterogeneity is partly explored in a subgroup analysis detailed in Section 6.4.4. Uncertainty regarding the appropriate source of data, and other assumptions are explored by scenario analysis and threshold analysis, as detailed in Section 6.4.3 and 6.4.5.

6.4.2 Base-case analysis

The base-case analysis considers two alternative set of comparisons. The first comparison is established between targeted software fusion and targeted cognitive fusion, while the second is established between combined software fusion and combined cognitive fusion. Therefore, we consider a dual base-case analysis with results presented separately for i) targeted biopsy alone and ii) combined (targeted + systematic) biopsy.

The dual base-case is defined by the following data sources and assumptions:

- The main analysis extension to the evidence synthesis for the subgroup of biopsy-naïve individuals, which uses:
  - the baseline distribution of test results for software fusion sourced from biopsy-naïve data from Filson et al. (2016).[96]
  - relative accuracy data from the multinomial evidence synthesis model (Model 1a) which was incorporated into the extension to the evidence synthesis - the network 1 was used to inform the targeted biopsy comparison, while network 2 informed the combined biopsy comparison;
  - Accuracy data from Mortezavi (2018) [107] extension to the evidence synthesis.
- The only differences between combined and targeted biopsy stem from the data used in the extension to the evidence synthesis (i.e., they are assumed to have the same profile of adverse events and biopsy procedure costs [note that both set of comparisons consider software fuion and cognitive fusion biopsy]);
- The cost of first and repeat biopsy with software fusion is modelled as an average of these costs for each technology (headline cost-effectiveness results for the individual technologies are presented for the base-case analysis);
- The cost of first biopsy assumed procedures are conducted as a mix of LATP and LATRUS; similarly, repeat biopsy is a mix of LATP, GATP and LATRUS. These proportions were assumed the same for software fusion and cognitive biopsy.
- Structural assumptions
  - Only individuals classified in the CPG 1 or ‘no cancer’ categories are eligible for repeat biopsy, and of those a fixed proportion received repeat biopsy in the model (15.45% and 5% for those classified CPGI and ‘no cancer, respectively);
  - While the model considers different progression rates by true CPG (as modified by radical treatment effect in accordance to the diagnosed CPG), progression across CPG scores is not modelled – only progression between each local disease status and the metastatic health state are possible);

– After two years in the misclassified localised disease health status, all individuals who remain in the corresponding states and have not yet received radical treatment will receive radical treatment according to their true disease status, incurring the costs and disutility of radical treatment then and receiving monitoring commensurate with their true disease status from that point onwards.

6.4.3 Threshold analysis on costs of software fusion

We have highlighted throughout Section 6.3.7 uncertainties and areas of evidence scarcity relating to the costs of the the biopsy procedure, particularly for the software fusion technologies. We reiterate that given these uncertainties and that it was not possible to calculate diagnostic performance evidence by individual software fusion devices at the granularity of classification (ISUP G1, ISUP G2, ISUP G3, and ISUP G 4 or 5) required by the economic model, it was considered the biopsy procedure costs for each individual technology were potentially misleading to decision makers.

Thus, we apply the average biopsy cost across all software fusion tehcnologies in this assessment for which cost data was submitted by the companies. We also perform a threshold analysis in which we estimate what is the cost per biopsy procedure with software fusion at which it is no longer likely that the new technologies will be cost-effective at the conventional range of opportunity costs considered by NICE. This threshold analysis applies the same assumptions and data sources of the base-case analysis, but assumes that:

- All biopsies are LATP procedures;
- Excludes the cost of the third-party ultrasounds from the biopsy cost calculations (to disentangle the cost of cognitive and software fusion).

These assumptions are necessary in order to run a threshold analysis varying a single parameter (i.e., cost of software fusion biopsy).

6.4.4 Subgroup analysis

As mentioned in Section 6.3.1, the extension of the evidence synthesis included a subgroup analysis for previous negative biopsy individuals. We performed a subgroup analysis for the same group of patients, which mirrors the subgroup analysis in Section 6.3.1. In brief, this subgroup analysis used the same evidence sources to inform the extension to the synthesis, except the baseline distribution of test results for software fusion. This was was sourced from previous negative biopsy data from Filson et al. (2016).[96]

6.4.5 Scenario analyses

The scenario analyses are summarised in Table 53. In brief, the aim of the scenario analysis is:

- Scenario analyses 1 and 2: to mirror the sensitivity analysis performed around the sources of data informing the sensitivity analyses of the evidence synthesis extension (see Section 6.3.1), and explore their impact on the cost-effectiveness estimates.
- Scenario analysis 3: to explore the impact of lowering the diagnostic accuracy of repeat biopsy, as considered in the PROMIS, NICE NG131 and Southampton DAR models.
- Scenario 4: to model the use of software fusion biopsy as quality assurance, as this was suggested by clinical advisers to the EAG as a potential value component of software. The clinical advisers commented that they would be more confident that a negative biopsy result with software fusion biopsy following a positive MRI result would not require a confirmatory biopsy compared to cognitive fusion, and that this confidence did not arise from any perceived gains in diagnostic accuracy of software fusion vs. cognitive fuion biopsy. Thus, we set the diagnostic accuracy of software fusion to be equal to that of the base- case accuracy for cognitive fusion (implying that the sole value of software fusion is to inform the selection of cases for repeat biopsy), and we changed the eligibility criteria for repeat biopsy with software fusion as described in Table 53.
- Scenario 5: aims to approximate the assumptions in regards to localised disease treatment conditional on final classification to those of the PROMIS model.
- Scenario 6.1 and 6.2: aim to explore the impact of using software fusion in NHS trusts with lower (6.1) and higher patient throughput (6.2) than that assumed to correspond in the base-case analysis to the national average throughput.

Table 53 Description of the scenario analyses

Scenario number and label	Element of uncertainty	Base-case	Scenario variation
1. PAIREDCAP (2019) baseline	Extension of the evidence synthesis model	Data sources for the extension to evidence synthesis: . baseline distribution of test results for software fusion from biopsy-naïve data in Filson et al. (2016) . Relative accuracy data from the multinomial evidence synthesis model (Model 1a, network 1+2 – targeted and combined biopsy) . Acccuracy data from Mortezavi (2018)	Data sources for the extension to evidence synthesis: . baseline distribution of test results for software fusion from biopsy-naïve data in biopsy-naïve data from PAIREDCAP (2019) for network 1. . Relative accuracy from the multinomial evidence synthesis model (Model 1a, network 1 only – targeted biopsy) . Acccuracy data as for base-case analysis.
2. Zhou (2018) diagnostic			Data sources for the extension to evidence synthesis: . baseline distribution of test results for software fusion as for base-case analysis.a . Relative accuracy from the multinomial evidence synthesis model (Model 1a, network 1 only – targeted biopsy) . Acccuracy data from Zhou et al.(2018)
3. Degradation of repeat biopsy accuracy	Diagnostic performance of MRI-influenced repeat biopsy	Repeat biopsy is as accurate as first biopsy for both cognitive and software fusion	. Probability of correctly classifying individuals as having cancer at each CPG category is reduced by 80% at repeat biopsies (changes in diagnostic accuracy are distributed equally across all other possible CPG classifications for each true disease CPG)
4. Software fusion as quality assurance	Diagnostic performance of MRI-influenced biopsy and selection for repeat biopsy	Diagnostic performance of MRI-influenced biopsy is informed by the extension of the evidence synthesis model 1a (network 1 and 2) and only a proportion of those classified at first biopsy as having no cancer or CPG1 receive repeat biopsy	. No difference in overall diagnostic performance of cognitive fusion vs. software fusion . Individuals eligible for repeat biopsy are those: ..Who have been misclassified as CPG1or no cancer at first biopsy with software fusion ..Who have been classified (correctly or not) as CPG1or no cancer at first biopsy with software fusion
5. Radical treatment for all identified CPG≥2 and conservative treatment for CPG 1	Distribution of treatment for localised disease	The distribution of radical treatment for localised disease is sourced from Parry et al. (2020)	. All individuals diagnosed CPG≥2 are treated at long-term model entrance with radical treatment (maintaining the distribution between radical prostatectomy and radiotherapy as per the base-case) and thise diagnosed CPG1 receive conservative treatment (and do not switch for radical treatment).
6.1 Throughput (150/year)	Annual biopsy throughput	300 biopsies per year	. 150 biopsies per year: 50% lower than base-case

6.2 Throughput (450/year)		300 biopsies per year	. 450 biopsies per year: 50% higher than base-case

6.4.6 Results

Base-case analysis

The deterministic and probabilistic cost-effectiveness results for the base-case analysis are presented in Table 54 and Table 55, respectively. The data sources used to derive prevalence and true disease status in this analysis refer to a biopsy naïve population. For both the targeted biopsy (informed by network 1 of Model 1a) and the combined biopsy (informed by network 2 of Model 1a) comparisons, the software fusion strategy seems to on average be costlier and to yield greater QALYs than the cognitive fusion strategy, resulting in a deterministic ICER of £5,623 and £1,826 per additional QALY, respectively. These ICERs are below the lower bound of the cost-effectiveness threshold range recommended by NICE, suggesting that it may be cost-effective compared to cognitive fusions in both the targetd and the combined comparisons. However, these results should be interpreted cautiously given the uncertainties in the relative diagnostic accuracy evidence which informs the model.

The probabilistic analysis suggests a higher probability of cost-effectiveness for software fusion vs. cognitive fusion at the range of cost-effectiveness thresholds recommended by NICE (0.64 and 0.68 at £20,000 and £30,000 per additional QALY for targeted software fusion biopsy). The probabilistic and deterministic cost-effectiveness results for each set of comparisons are similar. Henceforth and for subsequent analysis, we focus on the deterministic results, as these are easier to compare across base-case, threshold, subgroup and scenario analyses.

Table 54 Deterministic base-case cost-effectiveness results: i) targeted and ii) combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****
Targeted cognitive fusion	-0.00176	£445	11.45	8.29	£27,919	11.45	8.29	£28,364		6.87	7.34
Targeted software fusion	-0.00175	£543	11.46	8.30	£27,885	11.46	8.30	£28,428		6.88	7.35
Targeted	**Inc QALY loss**	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00001	£98	0.02	0.01	-£34	0.02	0.01	£63	£5,623	0.01	0.01
Strategy	**QALY loss**	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****
Combined cognitive fusion	-0.00177	£448	11.44	8.28	£27,889	11.44	8.28	£28,337		6.86	7.33
Combined software fusion	-0.00176	£544	11.49	8.31	£27,840	11.49	8.30	£28,384		6.89	7.36
Combined	**Inc QALY loss**	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00002	£95	0.05	0.03	-£49	0.05	0.03	£47	£1,826	0.02	0.02

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 55 Probabilistic base-case cost-effectiveness results: i) targeted and ii) combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results							
	QALY loss	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****	**Probability CE at £20,000****	**Probability CE at £30,000****
Targeted cognitive fusion	-0.00176	£445	11.46	8.30	£27,734	11.46	8.30	£28,179		6.89	7.36	0.36	0.32
Targeted software fusion	-0.00175	£543	11.48	8.31	£27,702	11.48	8.31	£28,245		6.90	7.37	0.64	0.68
Targeted	**Inc QALY loss**	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****		
Software fusion vs. cognitive fusion	0.00001	£98	0.02	0.01	-£32	0.02	0.01	£65	£6,197	0.01	0.01		
Strategy	**QALY loss**	**Total Costs**	**Total LYs***	**Total QALYs***	**Inc Costs***	**Inc LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****	**Probability CE at £20,000****	**Probability CE at £30,000****
Combined cognitive fusion	-0.00177	£448	11.46	8.30	£27,716	11.46	8.30	£28,164		6.89	7.36	0.27	0.25
Combined software fusion	-0.00176	£544	11.50	8.33	£27,669	11.50	8.32	£28,213		6.91	7.38	0.73	0.75
Combined	**Inc QALY loss**	**Inc Costs**	**Inc LYs***	**Total QALYs***	**Total Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****		
Software fusion vs. cognitive fusion	0.00002	£96	0.04	0.02	-£47	0.04	0.02	£49	£2,199	0.02	0.02		

*Discounted at 3.5% per annum; **Per additional QALY; CE, cost-effectiveness; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

For the targeted biopsy (network 1), the software fusion strategy results in average higher costs (£543 vs. £443) and slightly lower QALY loss due to biopsy adverse events (-0.00175 vs. 0.00176) compared to cognitive fusion in the diagnostic model. The higher costs are driven by the cost of performing biopsy with software fusion, which on average costs £92 and £97 more than with cognitive fusion, for first and repeat biopsy, respectively. The software fusion strategy appears to lead to fewer repeat biopsies due to its higher correct detection rate at categories CPG 2 to CPG 4-5 compared to cognitive fusion; this has a small impact on incremental costs and QALY loss. This small impact on costs and benefits is due to the reduction in repeat biopsy with software fusion compared to cognitive fusion being small (0.055 vs 0.050) and the only differences in rates of biopsy adverse events between MRI-influenced methods stemming from differences in the proportion of repeat biopsy for each strategy.

The targeted software fusion strategy appears to increase correct classification (see Table x, appendix) across all CPGs compared to targeted cognitive fusion at the end of the diagnostic pathway (final classification), particulary for CPG 2 (correctly classified 15% vs. 10%, out of a true disease prevalence for this category of 27%) and to a lesser extent for category CPG 1 (correctly classified 0.108 vs. 0.057, out of a true disease prevalence for this category of 32%). This is consistent with the results of the extension to the evidence synthesis (see Section 6.3.1) and suggests that even with repeat biopsy for the cases classified as CPG1 or no prostate cancer at first disease, the remaining true disease CPG 2 cases misclassified are likely to be largely classified as having no prostate cancer. For those with true disease CPG 3 (prevalence for this category of 18.3%) the increase in correct detection with software fusion vs. cognitive fusion is modest (from 9% to 9.5%). The likelihood of being CPG 3 and being misclassified by the cognitive fusion strategy as no cancer, CPG 1 or CPG 2 is 33%, 23% and 36%, respectively, whereas with software fusion these proportions are 40%, 10% and 39%, respectively (results not shown; extracted directly from model).

In the long-term model, the targeted software fusion strategy appears to be accompanied by small life-year and QALY gain (0.02 life years and 0.01 QALYs) compared to cognitive fusion in the long-term model. Some of the higher incremental diagnostic costs of the software fusion biopsy strategy vs. cognitive fusion appear to be offset by the lower costs accrued for this strategy compared to cognitive fusion in the long-term model (£28,885 vs £27,919). The higher health outcomes with the software fusion technology compared to cognitive fusion are likely to stem from a slight increase in time spent in the localised disease health state (as suggested by the higher life years and baseline QALYs accrued in the model and lower metastatic disease QALY loss; see Table x, Appendix 12), which is partially offset by the higher upfront QALY loss from immediate localised radical treatment with the software fusion strategy vs. cognitive fusion. This is due to more patients being correctly identified in

the diagnostic model with the software fusion strategy. The increased correct classification with software fusion also results in higher upfront costs from radical treatment and its adverse events, but lower costs of managing metastatic disease and of monitoring.

For combined biopsy (network 2), the incremental costs and QALY loss of the software fusion strategy vs. cognitive fusion in the diagnostic model are fairly similar to those observed for the targeted biopsy. However, there seems to be greater cost savings and health outcomes benefits in the long-term model for software fusion compared to cognitive fusion in the combined biopsy analysis, which result in cost-effectiveness resulst more favourable to the software fusion strategy.

The level of correct classification across all grades in the combined biopsy diagnostic pathway is increased for the software fusion strategy compared to cognitive fusion (45.5% vs. 68%, more so than for targeted biopsy). The results suggest that when compared to combined cognitive fusion strategy, software fusion retrieves a higher proportion of CPG 4-5 (8.5% vs. 3.4%), CPG 2 (20.7% vs. 8.9%) CPG 1 (15.4% vs. 9.6%). Overall, this suggests that 16.8% more individuals are correctly identified with combined software fusion vs. cognitive fusion at CPG 2 or above, the threshold above which radical treatment is a treatment option according to current clinical guidance.

The correct higher detection at CPG results warranting radical treatment results in higher costs of upfront radical treatment for combined software fusion compared to cognitive fusion, but also the health benefits in the long-term model (due to slower disease progression). It also reduces the costs of metastatic treatment for the software fusion strategy vs. cognitive fusion.The impact on total costs and QALYs in the long-term model is still limited, as the increased correct detection concentrates on those who have a true CPG 2 and are less likely to benefit from radical treatment than those at CPG 3-4 (where increases in correct classification for combined software fusion vs. cognitive fusion are less marked). Nevertheless, the cost savings (-£49) and small incremental increase on QALYs (0.03 QALYs) for combined software fusion compared to cognitive fusion, results in an ICER favourable to combined software fusion (see Table 54).

Results of base-case by software fusion technology

In Table 56 we show the deterministic base-case analysis analysis results of targeted software fusion by individual technology in pairwise comparison vs. targeted cognitive. Corresponding results for the combined comparison are presented in Appendix 12.

Table 56 Deterministic base-case cost-effectiveness results: targeted software fusion technologies pairwise comparisons with targeted cognitive fusion

Strategy	Diagnostic model	Overall results					
	Inc costs	Total LYs*	Total QALYs*	Total Costs*	ICER vs. cognitive fusion**	NHB at £20,000**	NHB at £30,000**
Targeted cognitive fusion	-	11.45	8.29	£28,364		6.87	7.34
Targeted software fusion	£98	11.46	8.30	£28,428	£5,623	6.88	7.35
Targeted bkFusion	£101			£28,431	£5,954	6.88	7.35
Targeted FusionVu	£125			£28,454	£8,001	6.88	7.35
Targeted Koelis Trinity	£105			£28,435	£6,302	6.88	7.35
Targeted Fusion Bx 2.0	£113			£28,443	£6,968	6.88	7.35
Targeted BiopSee	£44			£28,374	£890	6.88	7.35

*Discounted at 3.5% per annum; **Per additional QALY; CE, cost-effectiveness; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

The pairwise ICERs of the targeted software fusion strategies vs. cognitive fusion range between £28,374 and £28,454 per additional QALY for BiopSee and FusionVu, respectively. Results for the combined biopsy comparison show the same pattern. The only incremental difference between individual software fusion technologies strategies are the incremental costs in the diagnostic model.

Threshold analysis on costs of software fusion

Given the uncertainties in the costing of software fusion we conducted a threshold analysis to identify the software fusion biopsy cost at which there would be a shift in the decision to accept software fusion as a good use of NHS resources. Since the base-analysis suggests that the software fusion strategy might be cost-effective compared to cognitive fusion, the point of decision shift is identified as the cost per software fusion (holding the cost of cognitive fusion constant) at which the incremental of NHB of the software fusion biopsy compared to cognitive fusion becomes negative (i.e., software fusion is not likely to be cost-effective). The threshold analysis is conducted under the assumption that all biopsies are LATP and excluding the cost of the ultrasound components from the cost of cognitive fusion. Under these assumptions the cost per biopsy is £448.50 and £331.00 per software fusion and cognitive fusio biopsy, respectively.

The threshold analysis results (see Figure 18, Appendix 12) suggests that the decision inversion point is located at a cost per targeted software fusion biopsy of £586 and £695 at £20,000 and £30,000 per additional QALY, respectively. For combined software fusion biopsy, the inversion point cost per biopsy was estimated as of £874 and £1,116 at £20,000 and £30,000 per additional QALY, respectively.

Subgroup analysis

The deterministic cost-effectiveness results of the subgroup analysis for previous negative biopsy individuals are presented in Table 57. We note that this analysis only differs from the base-case analysis in the source for the baseline distribution of test results for software fusion (sourced from previous negative biopsy data from Filson et al. (2016)[96], rather than the biopsy naïve in the base-case analysis). The estimated prevalence of prostate cancer disease in this subgroup is lower than in the base-case analysis (57% vs. 88%), while the diagnostic accuracy matrices for both targeted and combined biopsies in the subgroup analysis (see Appendix 10) are similar to those estimated for biopsy naïve individuals (as expected).

In the subgroup analysis, there is an increased likelihood of correctly classifying individuals with prostate cancer across all CPGs for software vs. cognitive fusion in both the targeted and combined

biopsy analysis. However, the lower prevalence means that there are fewer individuals in the model who are more likely to benefit from radical treatment (e.g., prevalence at CPG 4-5 for the prior biopsy subgroup is 8.5% compared to 11.6% in the biopsy naïve). Consistently with this, the prior biopsy subgroup cost savings and QALY gains in the long-term model for software fusion vs. cognitive fusion strategies appear to be smaller than for the base-case (particularly so for combined biopsy strategies), resulting in increased ICERs compared to the biopsy naïve.

Table 57 Deterministic cost-effectiveness results for prior biopsy subgroup: i) targeted and ii) combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****
Targeted cognitive fusion	-0.00176	£444	11.75	8.68	£22,014	11.75	8.68	£22,457		7.56	7.93
Targeted software fusion	-0.00175	£542	11.76	8.69	£21,994	11.76	8.69	£22,536		7.56	7.94
Targeted	**Inc QALY loss**	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00000	£99	0.01	0.01	-£20	0.01	0.01	£79	£9,285	0.00	0.01
Strategy	**QALY loss**	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****
Combined cognitive fusion	-0.00177	£446	11.75	8.68	£22,001	11.75	8.68	£22,447		7.55	7.93
Combined software fusion	-0.00176	£545	11.77	8.69	£22,000	11.77	8.69	£22,545		7.57	7.94
Combined	**Inc QALY loss**	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00001	£98	0.03	0.02	-£1	0.03	0.02	£98	£5,946	0.01	0.01

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Scenario analysis

The summary results of the scenario analysis are presented in Table 58, with full breakdown presented in Appendix 12.

Table 58 Scenario analysis cost-effectiveness results: i) targeted and b) combined biopsy

Scenario	Inc LYs*	Inc QALYs*	Inc Costs*	ICER per QALY
Targeted biopsy				
Base-case	0.02	0.01	£63	£5,623
1. PAIREDCAP (2019) baseline	0.02	0.01	£39	£4,428
2. Zhou (2018) diagnostic	0.03	0.03	£83	£3,105
3. Degradation of repeat biopsy accuracy	0.02	0.01	£63	£5,477
4. Software fusion as quality assurance	0.000100	0.000099	£87	£875,042
5. Radical treatment for all identified CPG≥2	0.04	0.03	-£117	Dominates
6.1 Throughput (150/year)	0.02	0.01	£129	£11,425
6.2 Throughput (450/year)	0.02	0.01	£42	£3,689
Combined biopsy				
Base-case	0.04	0.02	£49	£2,199
1. PAIREDCAP (2019) baseline	-	-	-	-
2. Zhou (2018) diagnostic	-	-	-	-
3. Degradation of repeat biopsy accuracy	0.05	0.03	£46	£1,801
4. Software fusion as quality assurance	0.000141	0.000139	£81	£582,123
5. Radical treatment for all identified CPG≥2 and conservative treatment for CPG 1	0.08	0.05	-£300	0.08
6.1 Throughput (150/year)	0.05	0.03	£110	£4,275
6.2 Throughput (450/year)	0.05	0.03	£26	£1,009

*Software fusion compared to cognitive fusion; cost and health outcomes discounted at 3.5% per annum over the model time horizon

The cost-effectiveness results for both set of comparisons (targeted and combined biopsy) appear to be robust to variations of the elements of uncertainty in all scenario analyses, with the exception of scenario 5. We discuss below the scenarios in which data sources of the evidence synthesis extension were modified and scenario 5 given its high impact on the estimates of cost-effectiveness. The remaining scenarios are not discussed further.

In Scenario 1, the prevalence of prostate cancer is higher (at all CPGs except CPG1) than for the corresponding base-case analysis (targeted comparison), which means that there are proportionally more individuals who can potentially benefit from early treatment. The diagnostic accuracy of the targeted software fusion is also higher than that of cognitive fusion strategy, but more so to correctly identify those with CPG2. Overall, this translates into increased cost savings in the long-term model

for the targeted software fusion vs. cognitive fusion compared to the base-case (£-58 vs. -£34), which lead to a lower ICER.

In Scenario 2, the prevalence of prostate cancer is lower (at all CPGs except CPG2) than for the corresponding base-case analysis (targeted comparison), but the diagnostic accuracy is higher for software fusion compared to cognitive fusion for all categories of CPG, which overall reduces the ICER for the targeted software fusion strategy compared to cognitive fusion to ££3,689 per additional QALY.

Scenario 5 shows that if there is no difference in diagnostic accuracy between software fusion and cognitive fusion, even if some repeat biopsies can be avoided with software fusion due to it being less prone to operator inexperience, the ICERs for software fusion compared to cognitive fusion (targeted and combined biopsy analysis) are far above the upper bound of the cost-effectiveness threshold range recommended by NICE. This is because the small incremental benefits from fewer repeat biopsies are insufficient to offet the higher costs of software fusion biopsy compared to cognitive fusion.

7 DISCUSSION

7.1 Statement of principal findings

The systematic review of clinical evidence included a total of 3733 patients who received software fusion and 2154 individuals with cognitive fusion from 23 studies. Evidence was included for all devices specified in the protocol, except for Fusion Bx 2.0 and FusionVu. Fourteen studies were included in network meta-analyses.

Overall, the evidence for all devices was at high risk of bias and therefore the quantitative synthesis results must be interpreted with caution. Results from our main analysis (looking across ISUP grades) suggest that patients undergoing cognitive biopsy may show: i) a higher probability of being classified as not having cancer, ii) similar probability of being classified as having non-clinically significant cancer (ISUP grade 1), and iii) lower probability of being classified at higher ISUP grades, particularly ISUP 2. Similar results were obtained when comparing between same biopsy methods where both were combined with systematic biopsy.

Additional meta-analyses of cancer detection rates suggest that, compared with cognitive fusion biopsy, software fusion may identify more prostate cancer (any grade) (OR 1.30; 95% CrI 1.06, 1.61) and more non-clinically significant cancer (ISUP 1) (OR 1.98; 95% CrI 1.28, 3.06). Adding systematic biopsy to cognitive or software fusion may increase the detection of all prostate cancer and of clinically significant cancer, and from this evidence there is no suggestion that software fusion with concomitant systematic biopsy is superior to cognitive fusion with systematic biopsy.

Meta-analyses by individual device showed that compared with cognitive fusion biopsy, Biojet are Urostation are associated with a higher detection of prostate cancer overall. and that Biojet is associated with a higher rate of clinically significant cancers, although only one study of Biojet was included in the meta-analyses. Evidence for all other software devices was insufficient to reliably compare their accuracy with cognitive fusion, or to determine whether some software fusion technologies are more accurate than others. Evidence for bkFusion, ISR'obot Mona Lisa and Koelis Trinity was included in the systematic review but not in the meta-analyses. Compared with cognitive fusion, there was no evidence that the accuracy of software fusion may differ by lesion location, or between biopsy naïve and prior negative biopsy patients, or according to operator experience, although the number and quality of the studies informing the potential effect modifiers was limited.

Overall, there is no evidence that biopsy positivity rates and safety outcomes differ significantly between software fusion and cognitive fusion, or between software fusion devices. There was some evidence that systems with rigid registration (Biojet or Uronav) are easier and significantly faster to

use than elastic registration (KOELIS Trinity), although this is informed by a single, small study and is not conclusive.

The base-case cost-effectiveness analysis suggests for the targeted biopsy and the combined biopsy comparisons, that software fusion strategy is on average costlier and yields greater QALYs than the cognitive fusion strategy, resulting in a deterministic ICER of £5,623 and £1,826 per additional QALY for each comparison, respectively. These ICERs are below the lower bound of the cost-effectiveness threshold range recommended by NICE, suggesting that software fusion may be cost-effective compared to cognitive fusions in both the targetd and the combined comparisons. However, these results should be interpreted cautiously given the uncertainties in the relative diagnostic accuracy evidence which informs the model. The probabilistic analysis suggests a higher probability of cost-effectiveness for software fusion vs. cognitive fusion at the range of cost-effectiveness thresholds recommended by NICE (0.64 and 0.68 at £20,000 and £30,000 per additional QALY for targeted software fusion biopsy).

Given the uncertainties in the costing of software fusion we conducted a threshold analysis to identify the software fusion biopsy cost at which there would be a shift in the decision to accept software fusion as a good use of NHS resources. This suggested that at the cost of each the 5 individual technologies for which there was cost data, the recommendation decision would not change.

The base-case cost-effectiveness results were not sensitive to variations to alternative data sources and assumptions, except when no difference in diagnostic accuracy is assumed between software fusion and cognitive fusion. Under this assumption, the ICERs for software fusion compared to cognitive fusion (targeted and combined biopsy analysis) far exceed the upper bound of the cost-effectiveness threshold range recommended by NICE.

7.2 Strengths and limitations of the assessment

This is the first systematic review to formally compare the relative accuracy of software fusion and cognitive fusion, with and without systematic biopsy, as well as different software fusion devices, using both direct and indirect evidence in a formal network meta-analysis. In order to best estimate differences between biopsy methods for each prostate cancer grade, a multinomial logistic regression model was fitted, where the odds of being categorised in each of the different ISUP grades were allowed to vary by biopsy type.

Our findings are consistent with those of recent systematic reviews that found no statistically significant difference between software fusion and cognitive fusion at detecting clinically significant prostate cancers,[51, 53] although unlike recent evidence,[51-53] our network meta-analysis found that

software fusion increased detection of clinically insignificant cancer compared with cognitive fusion. This result might be explained by differences in review and synthesis methods.

Our review has a number of limitations. Despite attempts to reduce bias by excluding unpaired, non-randomised studies, the evidence included in the meta-analysis remains at high risk of bias. Although within-patient comparisons remove much of the risk of confounding from imbalances in participant characteristics, true blinding from tracks of preceding biopsy methods within the same examination is not feasible (or would require two separate biopsy sessions per patient, which would be unethical). So far, no high-quality RCTs have been published.

There was variation across the studies in patient characteristics. In particular, a number of studies included patients with prior negative biopsy and biopsy naïve patients, who form the large majority of patients eligible for targeted biopsy, were underrepresented. Some variation and gaps in reporting were observed in MRI acquisition methods, criteria for referral to biopsy, biopsy routes and anaesthesia methods. Definitions of prostate cancer and clinically significant cancer varied across the studies. There was insufficient evidence to explore the impact of a number of potential effect modifiers, including lesion location, operator experience, biopsy routes and anaesthesia methods.

Most estimates from the meta-analyses were imprecise, particularly in the multinomial synthesis and at higher ISUP grades where data was most sparse. The network meta-analysis relied on a number of assumptions. Cognitive fusion was assumed to be equivalent across studies. The risk and extent to which the accuracy of cognitive fusion may vary by centre and operator experience is uncertain due to lack of evidence. It was also assumed that data from within-patient studies were independent. A model that accounted for the full structure of the data was not available, although it could have added precision to the estimates.

There were few studies per comparison and not all studies reported outcomes by all cancer grades. Therefore, only fixed-effect models were fit to the data. Data was sparse for most software fusion devices, and few studies included more than one software fusion technology, making it difficult to draw conclusions for relative accuracy of individual devices. Many studies of software fusion devices were excluded as they were beyond the scope of this appraisal.

Whilst our review identified several relevant studies, many could not be included in the synthesis due to lack of reporting of key data. For example, studies comparing software and cognitive fusion to systematic biopsy reported data on both targeted technologies jointly, and few studies reported a sufficient breakdown of biopsy results by ISUP grades (or equivalent breakdown) to inform the evidence synthesis required for the economic model. In addition, where studies included a mixed population of patients, a lack of reporting of biopsy results for the relevant population led to their

exclusion from the meta-analysis. We were therefore limited in the models we could consider due to data sparseness, and results are uncertain.

Studies not included in the meta-analyses mostly reported test positive rates (positive cases as percentage of all patients). As this measure is dependent on disease prevalence rather than diagnostic accuracy, results from these studies may be influenced by differences in prostate cancer rates between cohorts and may not be reliable.

The above-mentioned limitations in the evidence are not captured in the quantitative evidence synthesis, which is used to inform the economic analysis.

The cost-effectiveness analysis relies on the evidence informing it. Beyond the evidence sourced from the synthesis, this includes evidence on the long-term outcomes of treating prostate cancer and the cost data on each software fusion technology. This evidence is limited.

7.3 Uncertainties

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

There was large uncertainty in all estimates due to the limited evidence. Meta-analyses showed moderate heterogeneity that could not be explained by differences in individual software fusion devices. The evidence for all software fusion devices was at high risk of bias, and the diagnostic accuracy of systematic biopsy relative to software fusion and cognitive may have been overestimated in the meta-analyses. The applicability of the evidence for KOELIS Trinity and Biopsee is uncertain. There is no evidence comparing the accuracy of Fusion Bx 2.0 and FusionVu with cognitive fusion, and no evidence for these devices were eligible for inclusion in the indirect comparisons.

None of the studies included in the systematic review of diagnostic accuracy used template mapping biopsy as a reference standard, and many studies did not use standard 12-core systematic biopsy in addition to targeted biopsy methods. This means that the absolute true rate of prostate cancer lesions was underestimated and is uncertain. However, the lack of a gold-standard test is likely to have affected comparisons between all devices similarly, and therefore is unlikely to have biased relative estimates of prostate cancer detection.

Where reported, the number of targeted cores performed with software and cognitive fusion were broadly comparable between the studies. However, not all studies reported data on number of targeted cores to fully assess the risk of confounding from a possible difference in number of targeted cores between software and cognitive fusion. Evidence for all other protocol specified outcomes was limited and inconclusive.

7.4 Other relevant factors

Participants of studies included in the systematic review of diagnostic accuracy and clinical effectiveness had elevated PSA and/or abnormal DRE results and were referred to targeted biopsy following a PI-RADS or Likert score of three or more on MRI. This is reflective of NICE guidance, which recommends that men should be referred for mpMRI if their PSA levels are above the age-specific reference range or if their prostate feels malignant on DRE. However, other organisations have recommended that PSA should be used as part of a risk prediction tool, potentially leading to better targeting of patients referred to mpMRI. It is unclear how a change in referral criteria may affect the applicability of this assessment's findings.[13]

8 CONCLUSIONS

8.1 Implications for service provision

Software fusion biopsies may identify more clinically insignificant cancer than cognitive fusion biopsies, although there is no evidence that software fusion detects more clinically significant cancer. Both software fusion and cognitive fusion biopsy miss clinically significant cancer lesions, and the addition of standard-systematic biopsy increases the detection of all prostate cancer and clinically significant cancer. There is insufficient evidence to conclude on the relative accuracy and clinical effectiveness of different software devices.

8.2 Suggested research priorities

High-quality, sufficiently powered randomised controlled trial evidence is needed to address the limitations of the diagnostic accuracy evidence identified in this assessment. IP7-PACIFIC (NCT05574647)[167], a large UK-based randomised trial, will aim to determine whether software fusion biopsy is superior to cognitive fusion at detecting clinically significant prostate cancers in patients with suspicious MRI in patients randomised to either mpMRI or bpMRI. Although in its early stages, it is hoped that this trial will provide more robust and precise diagnostic accuracy estimates.

Full reporting of ISUP grades for each randomised arm is recommended, and for within-patient comparison studies, full reporting of cross-tabulation tables, where the classification of patients' cancer by ISUP grade for each biopsy type is described and the relative accuracy of the interventions

can be derived. In mixed population studies, reporting by key patient characteristics, such as PI-RADS score, whether biopsy naive or experienced, and route of referral for MRI (e.g. following clinical concerns, routine surveillance, screening etc) are required to inform decision-making. Availability of more granular data from already published studies would enable future syntheses to make use of a larger body of evidence. Qualitative evidence on the acceptability of software fusion to patients, notably where biopsy procedure time might be significantly increased, is needed.

REFERENCES

1. Cancer Research UK. *Ten most common cancers in males*. 2021. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared#heading-One [accessed 20th June 2022].

2. Public Health England. *Cancer registration statistics, England: final release, 2018*. Public Health England; 2020. Available from: https://www.gov.uk/government/statistics/cancer-registration-statistics-england-2018-final-release/cancer-registration-statistics-england-final-release-2018 [accessed 2nd April 2022].

3. Cancer Research UK. *Prostate cancer statistics - Prostate cancer incidence*. 2017. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer#heading-Zero [accessed 23rd March 2022].

4. National Prostate Cancer Audit (NPCA) team. *Eighth year annual report of the NPCA prospective audit in England and Wales for men diagnosed from 1 April 2019 to 31 March 2020 and the impact of COVID-19 in England during 2020*. London: Royal College of Surgeons; 2022.

5. Jones AL, Chinegwundoh F. Update on prostate cancer in black men within the UK. *Ecancermedicalscience* 2014;**8**:455.

6. Kiciński M, Vangronsveld J, Nawrot TS. An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis. *PLoS One* 2011;**6**:e27130.

7. Guo Z, He J, Pan J, Huang L, Cao J, Bai Z, et al. Prevalence and risk factors for incidental prostate cancer in patients after transurethral resection of the prostate with negative results on prostate biopsy: a retrospective study. *Investig Clin Urol* 2022;**63**:201-6.

8. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;**344**:e3502.

9. McCaffery K, Nickel B, Pickles K, Moynihan R, Kramer B, Barratt A, et al. Resisting recommended treatment for prostate cancer: a qualitative analysis of the lived experience of possible overdiagnosis. *BMJ Open* 2019;**9**:e026960.

10. National Institute for Health and Care Excellence. *Prostate cancer: diagnosis and management. [I] Evidence reviews for risk stratification of localised prostate cancer. NICE guideline NG131*. London: NICE; 2021.

11. National Institute for Health and Care Excellence. *Diagnostics Assessment Programme. Diagnostics consultation document. Transperineal biopsy for diagnosing prostate cancer*. London: NICE; 2022.

12. National Institute of Health and Care Excellence. *Prostate cancer: diagnosis and management. NICE guideline [NG131] (Last updated: 15 December 2021)*. London: NICE; 2019. Available from: https://www.nice.org.uk/guidance/ng131

13. Prostate Cancer UK. *Consensus statements on PSA testing in asymptomatic men in the UK - information for health professionals*. Prostate Cancer UK; 2016. Available from: https://prostatecanceruk.org/PSAconsensusHP [accessed 24th October 2022].

14. National Institute for Health and Care Excellence. *Suspected cancer: recognition and referral [NG12]*. London: NICE; 2015.

15. NHS England. *Faster diagnosis standard*. NHS England; Available from: https://www.england.nhs.uk/cancer/faster-diagnosis/#:~:text=The%20Faster%20Diagnosis%20Standard%20(FDS,cancer%20have%20a%20timely%20diagnosis [accessed 24th October 2022].

16. National Institute for Health and Care Excellence. *Quality statement 1: Direct access to diagnostic tests. Suspected cancer - Quality standard [QS124]*. NICE; 2016. Available from: https://www.nice.org.uk/guidance/qs124/chapter/Quality-statement-1-Direct-access-to-diagnostic-tests [accessed 24th October 2022].

17. NHS England. *Implementing a timed prostate cancer diagnostic pathway. A handbook for local health and care systems*. London: NHS England; 2018.

18. Ahmed HU, Bosaily AES, Brown LC, Kaplan RS, Colaco-Moraes Y, Ward K, et al. The PROMIS study: a paired-cohort, blinded confirmatory study evaluating the accuracy of multi-parametric MRI and TRUS biopsy in men with an elevated PSA. *J Clin Oncol* 2016;**34**:5000.

19. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *N Engl J Med* 2018;**378**:1767-77.

20. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. *Eur Radiol* 2012;**22**:746-57.

21. Turkbey B, Choyke PL. Prostate magnetic resonance imaging: lesion detection and local staging. In: Klotman ME, editor. *Annual Review of Medicine, Vol 70*. Palo Alto: Annual Reviews; 2019. p. 451-9.

22. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* 2016;**69**:16-40.

23. West of Scotland Cancer Network. *Audit report. Prostate cancer quality performance indicators. Clinical audit data: 01 July 2019 to 30 June 2020*. Glasgow: West of Scotland Cancer Network; 2021.

24. North Cancer Alliance. *Quality performance indicators audit report. Prostate cancer*. Aberdeen: North Cancer Alliance; 2022.

25. Barrett T, Slough R, Sushentsev N, Shaida N, Koo BC, Caglic I, et al. Three-year experience of a dedicated prostate mpMRI pre-biopsy programme and effect on timed cancer diagnostic pathways. *Clin Radiol* 2019;**74**:894.e1-.e9.

26. Alabousi M, Salameh JP, Gusenbauer K, Samoilov L, Jafri A, Yu H, et al. Biparametric vs multiparametric prostate magnetic resonance imaging for the detection of prostate cancer in treatment-naïve patients: a diagnostic test accuracy systematic review and meta-analysis. *BJU Int* 2019;**124**:209-20.

27. Lopez JF, Campbell A, Omer A, Stroman L, Bondad J, Austin T, et al. Local anaesthetic transperineal (LATP) prostate biopsy using a probe-mounted transperineal access system: a multicentre prospective outcome analysis. *BJU Int* 2021;**128**:311-8.

28. National Institute for Health and Care Excellence. *NICE recommends new diagnostic devices for men with suspected prostate cancer in draft guidance*. NICE; 2022. Available from:

https://www.nice.org.uk/news/article/nice-recommends-new-diagnostic-devices-for-men-with-suspected-prostate-cancer-in-draft-guidance [accessed 9th May 2022].

29. National Institute for Health and Care Excellence. *Transperineal biopsy for diagnosing prostate cancer. In development [GID-DG10043]*. NICE; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-dg10043 [accessed 23rd May 2022].

30. Kucur M, Goktas S, Kaynar M, Apiliogullari S, Kilic O, Akand M, et al. Selective low-dose spinal anesthesia for transrectal prostate biopsy: a prospective and randomized study. *J Endourol* 2015;**29**:1412-7.

31. Wegelin O, Exterkate L, van der Leest M, Kummer JA, Vreuls W, de Bruin PC, et al. The FUTURE trial: a multicenter randomised controlled trial on target biopsy techniques based on magnetic resonance imaging in the diagnosis of prostate cancer in patients with prior negative biopsies. *Eur Urol* 2019;**75**:582-90.

32. Bjurlin MA, Taneja SS. Standards for prostate biopsy. *Curr Opin Urol* 2014;**24**:155-61.

33. Das CJ, Razik A, Sharma S, Verma S. Prostate biopsy: when and how to perform. *Clin Radiol* 2019;**74**:853-64.

34. European Association of Urology. *EAU guidelines on prostate cancer*. Arnhem, The Netherlands: EAU Guidelines Office 2022.

35. Fulgham PF, Rukstalis DB, Rubenstein JN, Taneja SS, Carroll PR, Pinto PA, et al. *Standard operating procedure for multiparametric magnetic resonance imaging in the diagnosis, staging and management of prostate cancer*. American Urological Association; 2019. Available from: https://www.auanet.org/guidelines-and-quality/guidelines/other-clinical-guidance/mri-of-the-prostate-sop [accessed 24th October 2022].

36. Prostate Cancer Foundation. *About prostate cancer*. Prostate Cancer Foundation; 2022. Available from: https://www.pcf.org/about-prostate-cancer/ [accessed 24th October 2022].

37. Rouviere O, Puech P, Renard-Penna R, Claudon M, Roy C, Mege-Lechevallier F, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. *Lancet Oncol* 2019;**20**:100-9.

38. van der Leest M, Cornel E, Israel B, Hendriks R, Padhani AR, Hoogenboom M, et al. Head-to-head comparison of transrectal ultrasound-guided prostate biopsy versus multiparametric prostate resonance imaging with subsequent magnetic resonance-guided biopsy in biopsy-naive men with elevated prostate-specific antigen: a large prospective multicenter clinical study. *Eur Urol* 2019;**75**:570-8.

39. Emmett L, Buteau J, Papa N, Moon D, Thompson J, Roberts MJ, et al. The additive diagnostic value of prostate-specific membrane antigen positron emission tomography computed tomography to multiparametric magnetic resonance imaging triage in the diagnosis of prostate cancer (PRIMARY): a prospective multicentre study. *Eur Urol* 2021;**80**:682-9.

40. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet* 2017;**389**:815-22.

41. Thompson J, Savdie R, Ponsky L, Brenner P, Shnier R, Moses D, et al. Magnetic resonance imaging detects significant prostate cancer and could be used to reduce unnecessary biopsies: initial results from a prospective trial of MRI in men planned for diagnostic biopsy. *Urology* 2013;**82**:S26-S7.

42. Parry MG, Boyle JM, Nossiter J, Morris M, Sujenthiran A, Berry B, et al. Determinants of variation in radical local treatment for men with high-risk localised or locally advanced prostate cancer in England [Published online ahead of print September 7 2021]. *Prostate Cancer Prostatic Dis* 2021:http://dx.doi.org/10.1038/s41391-021-00439-9.

43. Morris M. *Exploring the use of a new risk grouping to assess 'over-treatment' for the National Prostate Cancer Audit*. National Prostate Cancer Audit, Royal College of Surgeons of England; 2022. Available from: https://www.npca.org.uk/news/exploring-the-use-of-a-new-risk-grouping-to-assess-over-treatment-for-the-national-prostate-cancer-audit/ [accessed 22nd August 2022].

44. Bjurlin MA, Rosenkrantz AB, Taneja SS. MRI-fusion biopsy: the contemporary experience. *Transl Androl Uro* 2017;**6**:483-9.

45. Yan P, Wang XY, Huang W, Zhang Y. Local anesthesia for pain control during transrectal ultrasound-guided prostate biopsy: a systematic review and meta-analysis. *J Pain Res* 2016;**9**:787-96.

46. Lee D, Chung BH, Lee KS. Effect of training and individual operator's expertise on prostate cancer detection through prostate biopsy: implications for the current quantitative training evaluation system. *Investig Clin Urol* 2021;**62**:658-65.

47. Costa DN, Pedrosa I, Donato F, Jr., Roehrborn CG, Rofsky NM. MR imaging-transrectal US fusion for targeted prostate biopsies: implications for diagnosis and clinical management. *Radiographics* 2015;**35**:696-708.

48. Cash H, Gunzel K, Maxeiner A, Stephan C, Fischer T, Durmus T, et al. Prostate cancer detection on transrectal ultrasonography-guided random biopsy despite negative real-time magnetic resonance imaging/ultrasonography fusion-guided targeted biopsy: reasons for targeted biopsy failure. *BJU Int* 2016;**118**:35-43.

49. Al-Ahmadie HA, Tickoo SK, Olgac S, Gopalan A, Scardino PT, Reuter VE, et al. Anterior-predominant prostatic tumors: zone of origin and pathologic outcomes at radical prostatectomy. *Am J Surg Pathol* 2008;**32**:229-35.

50. Bott SRJ, Young MPA, Kellett MJ, Parkinson MC. Anterior prostate cancer: is it more difficult to diagnose? *BJU Int* 2002;**89**:886-9.

51. Watts KL, Frechette L, Muller B, Ilinksy D, Kovac E, Sankin A, et al. Systematic review and meta-analysis comparing cognitive vs. image-guided fusion prostate biopsy for the detection of prostate cancer. *Urol Oncol* 2020;**38**:734.e19-.e25.

52. Sathianathen NJ, Butaney M, Bongiorno C, Konety BR, Bolton DM, Lawrentschuk N. Accuracy of the magnetic resonance imaging pathway in the detection of prostate cancer: a systematic review and meta-analysis. *Prostate Cancer Prostatic Dis* 2019;**22**:39-48.

53. Bass EJ, Pantovic A, Connor MJ, Loeb S, Rastinehad AR, Winkler M, et al. Diagnostic accuracy of magnetic resonance imaging targeted biopsy techniques compared to transrectal ultrasound guided biopsy of the prostate: a systematic review and meta-analysis. *Prostate Cancer Prostatic Dis* 2021;**21**:21.

54. Valerio M, Donaldson I, Emberton M, Ehdaie B, Hadaschik BA, Marks LS, et al. Detection of clinically significant prostate cancer using magnetic resonance imaging-ultrasound fusion targeted biopsy: a systematic review. *Eur Urol* 2015;**68**:8-19.

55. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. *N Engl J Med* 2016;**375**:1415-24.

56. Kishan AU, Sun Y, Hartman H, Pisansky TM, Bolla M, Neven A, et al. Androgen deprivation therapy use and duration with definitive radiotherapy for localised prostate cancer: an individual patient data meta-analysis. *Lancet Oncol* 2022;**23**:304-16.

57. National Institute for Health and Care Excellence. *Low dose rate brachytherapy for localised prostate cancer (IPG 132)*. London: NICE; 2005.

58. Morris WJ, Tyldesley S, Rodda S, Halperin R, Pai H, McKenzie M, et al. Androgen suppression combined with elective nodal and dose escalated radiation therapy (the ASCENDE-RT trial): an analysis of survival endpoints for a randomized trial comparing a low-dose-rate brachytherapy boost to a dose-escalated external beam boost for high- and intermediate-risk prostate cancer. *Int J Radiat Oncol Biol Phys* 2017;**98**:275-85.

59. James ND, Sydes MR, Clarke NW, Mason MD, Dearnaley DP, Spears MR, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet* 2016;**387**:1163-77.

60. Fizazi K, Faivre L, Lesaunier F, Delva R, Gravis G, Rolland F, et al. Androgen deprivation therapy plus docetaxel and estramustine versus androgen deprivation therapy alone for high-risk localised prostate cancer (GETUG 12): a phase 3 randomised controlled trial. *Lancet Oncol* 2015;**16**:787-94.

61. Schweizer MT, Huang P, Kattan MW, Kibel AS, de Wit R, Sternberg CN, et al. Adjuvant leuprolide with or without docetaxel in patients with high-risk prostate cancer after radical prostatectomy (TAX-3501): important lessons for future trials. *Cancer* 2013;**119**:3610-8.

62. Wegelin O, van Melick HHE, Hooft L, Bosch J, Reitsma HB, Barentsz JO, et al. Comparing Three Different Techniques for Magnetic Resonance Imaging-targeted Prostate Biopsies: A Systematic Review of In-bore versus Magnetic Resonance Imaging-transrectal Ultrasound fusion versus Cognitive Registration. Is There a Preferred Technique? *Eur Urol* 2017;**71**:517-31.

63. Drost F, Osses D, Nieboer D, Steyerberg E, Bangma C, Roobol M, et al. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. *Cochrane Database of Systematic Reviews* 2019;**4**.

64. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol* 2016;**75**:40-6.

65. Beltran L, Ahmad AS, Sandu H, Kudahetti S, Soosay G, Møller H, et al. Histopathologic false-positive diagnoses of prostate cancer in the age of immunohistochemistry. *Am J Surg Pathol* 2019;**43**:361-8.

66. Drost FJ, Osses D, Nieboer D, Bangma C, Steyerberg E, Roobol M, et al. Prostate MRI, with or without targeted biopsy, and standard biopsy for detecting prostate cancer: A cochrane systematic review and meta-analysis. *J Urol* 2019;**201(4 Supplement 1)**:e1097.

67. 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.

68. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF, et al. QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. *Ann Intern Med* 2021;**174**:1592-9.

69. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU Technical Support Document 1: Introduction to evidence synthesis for decision making*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2011.

70. Dias S, Caldwell DM. Network meta-analysis explained. *Arch Dis Child Fetal Neonatal Ed* 2019;**104**:F8-F12.

71. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials* Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2011.

72. R Core Team. *R: a language and environment for statistical computing [software]*. Foundation for Statistical Computing; 2022. Available from: https://www.R-project.org/. [accessed 25th October 2022].

73. Rücker G, Schwarzer G, Krahn U, König J. *netmeta: Network meta-analysis using frequentist methods. version 2.5-0 ed. [software]*. 2022. Available from: http://CRAN.R-project.org/package=netmeta [accessed 25th October 2022].

74. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*. 3rd ed. Boca Raton, FL: CRC Press; 2013.

75. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. *The BUGS book*. Boca Raton, FL: CRC Press; 2013.

76. van Valkenhoef G, Kuiper J. *gemtc: Network neta-analysis using Bayesian methods. R package. 1.0-1 ed. [software]*. CRAN; 2021. Available from: https://CRAN.R-project.org/package=gemtc [accessed 25th October 2022].

77. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. *NICE DSU Technical Support Document 4: Inconsistency in networks of evidence based on randomised controlled trials* Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2011.

78. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* 2020;**368**:l6890.

79. Thangarasu M, Jayaprakash SP, Selvaraj N, Bafna S, Paul R, Mahesh C, et al. A prospective study on the efficacy of cognitive targeted transrectal ultrasound prostate biopsy in diagnosing clinically significant prostate cancer. *Res Rep Urol* 2021;**13**:207-13.

80. Alberts AR, Schoots IG, Bokhorst LP, Drost FH, van Leenders GJ, Krestin GP, et al. Characteristics of prostate cancer found at fifth screening in the European randomized study of screening for prostate cancer Rotterdam: can we selectively detect high-grade prostate cancer with upfront multivariable risk stratification and magnetic resonance imaging? *Eur Urol* 2018;**73**:343-50.

81. Ferriero M, Tuderti G, Muto GL, Fiori C, Bove AM, Mastroianni R, et al. Diagnostic performance of fusion (US/MRI guided) prostate biopsy: propensity score matched comparison of elastic versus rigid fusion system. *World J Urol* 2022;**40**:991-6.

82. Izadpanahi MH, Elahian A, Gholipour F, Khorrami MH, Zargham M, Mohammadi Sichani M, et al. Diagnostic yield of fusion magnetic resonance-guided prostate biopsy versus cognitive-guided biopsy in biopsy-naive patients: a head-to-head randomized controlled trial. *Prostate Cancer Prostatic Dis* 2021;**24**:1103-9.

83. Sokolakis I, Pyrgidis N, Koneval L, Krebs M, Thurner A, Kubler H, et al. Usability and diagnostic accuracy of different MRI/ultrasound-guided fusion biopsy systems for the detection of clinically significant and insignificant prostate cancer: a prospective cohort study. *World J Urol* 2021;**39**:4101-8.

84. Rabah D, Al-Taweel W, Khan F, Arafa M, Mehmood S, Mokhtar A, et al. Transperineal versus transrectal multi-parametric magnetic resonance imaging fusion targeted prostate biopsy. *Saudi Med J* 2021;**42**:649-54.

85. Liang L, Cheng Y, Qi F, Zhang L, Cao D, Cheng G, et al. A comparative study of prostate cancer detection rate between transperineal COG-TB and transperineal FUS-TB in patients with PSA <=20 ng/mL. *J Endourol* 2020;**34**:1008-14.

86. Kulis T, Zekulic T, Alduk AM, Lusic M, Bulimbasic S, Ferencak V, et al. Targeted prostate biopsy using a cognitive fusion of multiparametric magnetic resonance imaging and transrectal ultrasound in patients with previously negative systematic biopsies and non-suspicious digital rectal exam. *Croat Med J* 2020;**61**:49-54.

87. Wajswol E, Winoker JS, Anastos H, Falagario U, Okhawere K, Martini A, et al. A cohort of transperineal electromagnetically tracked magnetic resonance imaging/ultrasonography fusion-guided biopsy: assessing the impact of inter-reader variability on cancer detection. *BJU Int* 2020;**125**:531-40.

88. Elkhoury FF, Felker ER, Kwan L, Sisk AE, Delfin M, Natarajan S, et al. Comparison of targeted vs systematic prostate biopsy in men who are biopsy naive: the prospective assessment of image registration in the diagnosis of prostate cancer (PAIREDCAP) study. *JAMA Surg* 2019;**154**:811-8.

89. Stabile A, Dell'Oglio P, Gandaglia G, Fossati N, Brembilla G, Cristel G, et al. Not all multiparametric magnetic resonance imaging-targeted biopsies are equal: the impact of the type of approach and operator expertise on the detection of clinically significant prostate cancer. *Eur Urol Oncol* 2018;**1**:120-8.

90. Monda SM, Vetter JM, Andriole GL, Fowler KJ, Shetty AS, Weese JR, et al. Cognitive versus software fusion for MRI-targeted biopsy: experience before and after implementation of fusion. *Urology* 2018;**119**:115-20.

91. Kaufmann S, Russo GI, Thaiss W, Notohamiprodjo M, Bamberg F, Bedke J, et al. Cognitive versus software-assisted registration: development of a new nomogram predicting prostate cancer at MRI-targeted biopsies. *Clin Genitourin Cancer* 2018;**16**:e953-e60.

92. Fourcade A, Payrard C, Tissot V, Perrouin-Verbe MA, Demany N, Serey-Effeil S, et al. The combination of targeted and systematic prostate biopsies is the best protocol for the detection of clinically significant prostate cancer. *Scand J Urol* 2018;**52**:174-9.

93. Cornud F, Roumiguie M, de Longchamps NB, Ploussard G, Bruguiere E, Portalez D, et al. Precision matters in MR imaging-targeted prostate biopsies: Evidence from a prospective study of cognitive and elastic fusion registration transrectal biopsies. *Radiology* 2018;**287**:534-42.

94. Albisinni S, Aoun F, Noel A, El Rassy E, Lemort M, Paesmans M, et al. Are concurrent systematic cores needed at the time of targeted biopsy in patients with prior negative prostate biopsies? *Prog Urol* 2018;**28**:18-24.

95. Hansen NL, Barrett T, Kesch C, Pepdjonovic L, Bonekamp D, O'Sullivan R, et al. Multicentre evaluation of magnetic resonance imaging supported transperineal prostate biopsy in biopsy-naive men with suspicion of prostate cancer. *BJU Int* 2018;**122**:40-9.

96. Filson CP, Natarajan S, Margolis DJ, Huang J, Lieu P, Dorey FJ, et al. Prostate cancer detection with magnetic resonance-ultrasound fusion biopsy: the role of systematic and targeted biopsies. *Cancer* 2016;**122**:884-92.

97. Wysock JS, Rosenkrantz AB, Huang WC, Stifelman MD, Lepor H, Deng FM, et al. A prospective, blinded comparison of magnetic resonance (MR) imaging-ultrasound fusion and visual estimation in the performance of MR-targeted prostate biopsy: the PROFUS trial. *Eur Urol* 2014;**66**:343-51.

98. Delongchamps NB, Peyromaure M, Schull A, Beuvon F, Bouazza N, Flam T, et al. Prebiopsy magnetic resonance imaging and prostate cancer detection: comparison of random and targeted biopsies. *J Urol* 2013;**189**:493-9.

99. Gomez-Ortiz D, Garza-Gangemi AM, Oropeza-Aguilar M, Rangel-Suarez S, Espinosa-Cruz V, Villegas-Hernandez AC, et al. Routine systematic prostate biopsies not replaced by magnetic resonance imaging-targeted biopsy [published online ahead of print June 7 2022]. *Rev Invest Clin* 2022:https://dx.doi.org/10.24875/RIC.22000084.

100. Lockhart K, Martin J, White M, Raman A, Grant A, Chong P. Fusion versus cognitive MRI-guided prostate biopsies in diagnosing clinically significant prostate cancer [published online ahead of print May 13 2022]. *J Clin Urol* 2022:https://dx.doi.org/10.1177/20514158221085081.

101. Kaufmann S, Russo GI, Bamberg F, Lowe L, Morgia G, Nikolaou K, et al. Prostate cancer detection in patients with prior negative biopsy undergoing cognitive-, robotic- or in-bore MRI target biopsy. *World J Urol* 2018;**36**:761-8.

102. Exterkate L, Wegelin O, Barentsz JO, van der Leest MG, Kummer JA, Vreuls W, et al. Is there still a need for repeated systematic biopsies in patients with previous negative biopsies in the era of magnetic resonance imaging-targeted biopsies of the prostate? *Eur Urol Oncol* 2020;**3**:216-23.

103. Kalavagunta C, Zhou X, Schmechel SC, Metzger GJ. Registration of in vivo prostate MRI and pseudo-whole mount histology using Local Affine Transformations guided by Internal Structures (LATIS). *J Magn Reson Imaging* 2015;**41**:1104-14.

104. Wegelin O, Exterkate L, van der Leest M, Kelder JC, Ruud Bosch JLH, Barentsz JO, et al. Complications and adverse events of three magnetic resonance imaging-based target biopsy techniques in the diagnosis of prostate cancer among men with prior negative biopsies: results from the FUTURE trial, a multicentre randomised controlled trial. *Eur Urol Oncol* 2019;**2**:617-24.

105. Jordan PWTBMIL, Weerdmeester B, Brooke J. SUS: a 'quick and dirty' usability scale. In: *Usability Evaluation In Industry*. London: CRC Press; 1996. p. 189–94.

106. Mannaerts CK, Engelbrecht MRW, Postema AW, van Kollenburg RAA, Hoeks CMA, Savci-Heijink CD, et al. Detection of clinically significant prostate cancer in biopsy-naive men: direct comparison of systematic biopsy, multiparametric MRI- and contrast-ultrasound-dispersion imaging-targeted biopsy. *BJU Int* 2020;**126**:481-93.

107. Mortezavi A, Marzendorfer O, Donati OF, Rizzi G, Rupp NJ, Wettstein MS, et al. Diagnostic accuracy of multiparametric magnetic resonance imaging and fusion guided targeted biopsy evaluated by transperineal template saturation prostate biopsy for the detection and characterization of prostate cancer. *J Urol* 2018;**200**:309-18.

108. Zhou Y, Zhou Z, Li Q, Xu Y, Sun H, Xiao Y, et al. Diagnostic accuracy of magnetic resonance-guided prostate biopsy and template-guided transperineal saturation biopsy. *Medicine* 2018;**97**:e12495.

109. Simmons LAM, Kanthabalan A, Arya M, Briggs T, Barratt D, Charman SC, et al. Accuracy of transperineal targeted prostate biopsies, visual estimation and image fusion in men needing repeat biopsy in the PICTURE trial. *J Urol* 2018;**200**:1227-34.

110. Hansen N, Patruno G, Wadhwa K, Gaziev G, Miano R, Barrett T, et al. Magnetic Resonance and Ultrasound Image Fusion Supported Transperineal Prostate Biopsy Using the Ginsburg Protocol: Technique, Learning Points, and Biopsy Results. *Eur Urol* 2016;**70**:332-40.

111. Kesch C, Radtke JP, Popeneciu IV, Gasch C, Dieffenbacher SC, Klein T, et al. TOP: Prospective evaluation of a volume based, computer assisted method for transperineal optimized prostate biopsy. *Urol Int* 2017;**99**:149-55.

112. Hansen NL, Caglic I, Berman LH, Kastner C, Doble A, Barrett T. Multiparametric Prostate Magnetic Resonance Imaging and Cognitively Targeted Transperineal Biopsy in Patients With Previous Abdominoperineal Resection and Suspicion of Prostate Cancer. *Urology* 2016;**96**:8-14.

113. Rodda S, Tyldesley S, Morris WJ, Keyes M, Halperin R, Pai H, et al. ASCENDE-RT: an analysis of treatment-related morbidity for a randomized trial comparing a low-dose-rate brachytherapy boost with a dose-escalated external beam boost for high- and intermediate-risk prostate cancer. *Int J Radiat Oncol Biol Phys* 2017;**98**:286-95.

114. Incrocci L, Wortel RC, Alemayehu WG, Aluwini S, Schimmel E, Krol S, et al. Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer (HYPRO): final efficacy results from a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol* 2016;**17**:1061-9.

115. Catton CN, Lukka H, Gu CS, Martin JM, Supiot S, Chung PWM, et al. Randomized trial of a hypofractionated radiation regimen for the treatment of localized prostate cancer. *J Clin Oncol* 2017;**35**:1884-90.

116. Dearnaley D, Syndikus I, Mossop H, Khoo V, Birtle A, Bloomfield D, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. *Lancet Oncol* 2016;**17**:1047-60.

117. Widmark A, Gunnlaugsson A, Beckman L, Thellenberg-Karlsson C, Hoyer M, Lagerlund M, et al. Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial. *Lancet* 2019;**394**:385-95.

118. Marzi S, Saracino B, Petrongari MG, Arcangeli S, Gomellini S, Arcangeli G, et al. Modeling of alpha/beta for late rectal toxicity from a randomized phase II study: conventional versus hypofractionated scheme for localized prostate cancer. *J Exp Clin Cancer Res* 2009;**28**:117.

119. Wilt TJ, Brawer MK, Jones KM, Barry MJ, Aronson WJ, Fox S, et al. Radical prostatectomy versus observation for localized prostate cancer. *N Engl J Med* 2012;**367**:203-13.

120. Bill-Axelson A, Holmberg L, Ruutu M, Garmo H, Stark JR, Busch C, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med* 2011;**364**:1708-17.

121. Hegde JV, Demanes DJ, Veruttipong D, Raince J, Park SJ, Raman SS, et al. Pretreatment 3T multiparametric MRI staging predicts for biochemical failure in high-risk prostate cancer treated with combination high-dose-rate brachytherapy and external beam radiotherapy. *Brachytherapy* 2017;**16**:1106-12.

122. Gnanapragasam VJ, Lophatananon A, Wright KA, Muir KR, Gavin A, Greenberg DC. Improving clinical risk stratification at diagnosis in primary prostate cancer: a prognostic modelling study. *PLoS Med* 2016;**13**:e1002063.

123. James ND, Spears MR, Clarke NW, Dearnaley DP, De Bono JS, Gale J, et al. Survival with newly diagnosed metastatic prostate cancer in the "docetaxel era": data from 917 patients in the control arm of the STAMPEDE Trial (MRC PR08, CRUK/06/019). *Eur Urol* 2015;**67**:1028-38.

124. Bryant RJ, Oxley J, Young GJ, Lane JA, Metcalfe C, Davis M, et al. The ProtecT trial: analysis of the patient cohort, baseline risk stratification and disease progression. *BJU Int* 2020;**125**:506-14.

125. Yang Y, Abel L, Buchanan J, Fanshawe T, Shinkins B. Use of decision modelling in economic evaluations of diagnostic tests: an appraisal and review of Health Technology Assessments in the UK. *Pharmacoecon Open* 2019;**3**:281-91.

126. Souto-Ribeiro I, Woods L, Maund E, Scott DA, Lord J, Picot J, et al. *Transperineal biopsy in people with suspected prostate cancer - a systematic review and economic evaluation*. London: NICE; 2022.

127. Pahwa S, Schiltz NK, Ponsky LE, Lu Z, Griswold MA, Gulani V. Cost-effectiveness of MR Imaging-guided Strategies for Detection of Prostate Cancer in Biopsy-Naive Men. *Radiology* 2017;**285**:157-66.

128. Hayes JH, Ollendorf DA, Pearson SD, Barry MJ, Kantoff PW, Lee PA, et al. Observation versus initial treatment for men with localized, low-risk prostate cancer: a cost-effectiveness analysis. *Ann Intern Med* 2013;**158**:853-60.

129. Cheng LJ, Soon SS, Tan TW, Tan CH, Lim TSK, Tay KJ, et al. Cost-effectiveness of MRI targeted biopsy strategies for diagnosing prostate cancer in Singapore. *BMC Health Serv Res* 2021;**21**:909.

130. Hao S, Karlsson A, Heintz E, Elfstrom KM, Nordstrom T, Clements M. Cost-Effectiveness of Magnetic Resonance Imaging in Prostate Cancer Screening: A Microsimulation Study. *Value Health* 2021;**24**:1763-72.

131. Wilson ECF, Wreford A, Tamer P, Leonard K, Brechka H, Gnanapragasam VJ. Economic evaluation of transperineal versus transrectal devices for local anaesthetic prostate biopsies. *Pharmacoeconomics* 2021;**5**:737-53.

132. Getaneh AM, Heijnsdijk EA, de Koning HJ. Cost-effectiveness of multiparametric magnetic resonance imaging and MRI-guided biopsy in a population-based prostate cancer screening setting using a micro-simulation model. *Cancer Medicine* 2021;**10**:4046-53.

133. National Institute for Health and Care Excellence. *NG131 Health economic model report - Prostate cancer update*. London: NICE; 2019.

134. Barnett CL, Davenport MS, Montgomery JS, Wei JT, Montie JE, Denton BT. Cost-effectiveness of magnetic resonance imaging and targeted fusion biopsy for early detection of prostate cancer. *BJU Int* 2018;**122**:50-8.

135. Faria R, Soares MO, Spackman E, Ahmed HU, Brown LC, Kaplan R, et al. Optimising the Diagnosis of Prostate Cancer in the Era of Multiparametric Magnetic Resonance Imaging: A Cost-effectiveness Analysis Based on the Prostate MR Imaging Study (PROMIS). *Eur Urol* 2018;**73**:23-30.

136. Brown LC, Ahmed HU, Faria R, El-Shater Bosaily A, Gabe R, Kaplan RS, et al. Multiparametric MRI to improve detection of prostate cancer compared with transrectal ultrasound-guided prostate biopsy alone: the PROMIS study. *Health Technol Assess* 2018;**22**:1-176.

137. Patel S, Rongen JJ, Futterer JJ, Boltyenkov A, Rovers MM. The Role of Multiparametric Magnetic Resonance Imaging in Active Surveillance for Men with Low-risk Prostate Cancer: A Cost-effectiveness Modeling Study. *Eur Urol Oncol* 2018;**1**:476-83.

138. Sathianathen NJ, Kuntz KM, Alarid-Escudero F, Lawrentschuk NL, Bolton DM, Murphy DG, et al. Incorporating biomarkers into the primary prostate biopsy setting: a cost-effectiveness analysis. *J Urol* 2018;**200**:1215-20.

139. Pahwa S, Schiltz NK, Ponsky LE, Lu Z, Griswold MA, Gulani V. MR Imaging-Guided strategies for detection of prostate cancer in biopsy-naive men: Response. *Radiology* 2017;**285(3)**:1053.

140. Venderink W, Govers TM, de Rooij M, Futterer JJ, Sedelaar JPM. Cost-Effectiveness Comparison of Imaging-Guided Prostate Biopsy Techniques: Systematic Transrectal Ultrasound, Direct In-Bore MRI, and Image Fusion. *AJR Am J Roentgenol* 2017;**208**:1058-63.

141. Cerantola Y, Dragomir A, Tanguay S, Bladou F, Aprikian A, Kassouf W. Cost-effectiveness of multiparametric magnetic resonance imaging and targeted biopsy in diagnosing prostate cancer. *Urol Oncol* 2016;**34**:119.e1-9.

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

143. Mowatt G, Scotland G, Boachie C, Cruickshank M, Ford JA, Fraser C, et al. The diagnostic accuracy and cost-effectiveness of magnetic resonance spectroscopy and enhanced magnetic resonance imaging techniques in aiding the localisation of prostate abnormalities for biopsy: a systematic review and economic evaluation. *Health Technol Assess* 2013;**17**:vii-xix, 1-281.

144. Royal College of Surgeons of England. *NPCA Annual Report 2020*. London: Royal College of Surgeons of England; 2021.

145. Bill-Axelson A, Holmberg L, Garmo H, Rider JR, Taari K, Busch C, et al. Radical prostatectomy or watchful waiting in early prostate cancer. *N Engl J Med* 2014;**370**:932-42.

146. Torvinen S, Färkkilä N, Sintonen H, Saarto T, Roine RP, Taari K. Health-related quality of life in prostate cancer. *Acta Oncol* 2013;**52**:1094-101.

147. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;**13**:509-18.

148. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. *N Engl J Med* 2016;**375**:1425-37.

149. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: a modelling study. *Palliat Med* 2015;**29**:899-907.

150. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000;**10**:325-37.

151. Clarke NW, Ali A, Ingleby FC, Hoyle A, Amos CL, Attard G, et al. Addition of docetaxel to hormonal therapy in low- and high-burden metastatic hormone sensitive prostate cancer: long-term survival results from the STAMPEDE trial. *Ann Oncol* 2019;**30**:1992-2003.

152. *National life tables: UK*. Office for National Statistics; 2021. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables [accessed 29th September 2022].

153. Ankit Rohatgi. *WebPlotDigitizer*. Ankit Rohatgi; 2022. Available from: https://automeris.io/WebPlotDigitizer/index.html [accessed 27th October 2022].

154. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9.

155. Jackson C. flexsurv: a platform for parametric survival modeling in R. *J Stat Softw* 2016;**70**:1 - 33.

156. Brent R. *Algorithms for minimization without derivatives*. Mineloa: Dover Publications; 2022.

157. Wallis CJ. *ESMO 2021: Final overall survival analysis from ARCHES: a phase 3, randomized, double-blind, placebo-controlled study of enzalutamide + ADT in men with mHSPC*. UroToday.com; 2021. Available from: https://www.urotoday.com/conference-highlights/esmo-2021/esmo-2021-prostate-cancer/132209-esmo-2021-lba25-final-overall-survival-os-analysis-from-arches-a-phase-3-randomized-double-blind-placebo-pbo-controlled-study-of-enzalutamide-enza-androgen-deprivation-therapy-adt-in-men-with-metastatic-hormone-sensitive-prostate-cancer-mhspc.html [accessed 13th October 2022].

158. Chi KN, Chowdhury S, Bjartell A, Chung BH, Pereira de Santana Gomes AJ, Given R, et al. Apalutamide in patients with metastatic castration-sensitive prostate cancer: Final survival analysis of the randomized, double-blind, phase III TITAN study. *J Clin Oncol* 2021;**39**:2294-303.

159. Parry MG, Cowling TE, Sujenthiran A, Nossiter J, Berry B, Cathcart P, et al. Risk stratification for prostate cancer management: value of the Cambridge Prognostic Group classification for assessing treatment allocation. *BMC Med* 2020;**18**:114.

160. Souto-Ribeiro I, Woods L, Maund E, Scott DA, Lord J, Picot J, et al. *Transperineal biopsy in people with suspected prostate cancer - a systematic review and economic evaluation. Addendum 1*. London: NICE; 2022.

161. Armstrong AJ, Iguchi T, Azad AA, Szmulewitz RZ, Holzbeierlein J, Villers A, et al. LBA25 Final overall survival (OS) analysis from ARCHES: a phase III, randomized, double-blind, placebo (PBO)-controlled study of enzalutamide (ENZA) + androgen deprivation therapy (ADT) in men with metastatic hormone-sensitive prostate cancer (mHSPC). *Ann Oncol* 2021;**32**:S1300-1.

162. Exchange Rates UK. *US Dollar (USD) to British Pound (GBP) exchange rate history. Wednesday 9/03/2022 to Saturday 3/09/2022*. Exchange Rates UK; 2022. Available from: https://www.exchangerates.org.uk/USD-GBP-exchange-rate-history.html [accessed 5th September 2022].

163. Jones K, Burns A. *Unit costs of health and social care 2021*. Canterbury: Personal Social Services Research Unit, University of Kent; 2021.

164. Cash HH, S L Shore, N Pavlovich, C P Bulang, S Schostak, M Planken, E Prostate cancer detection by novice micro-ultrasound users enrolled in a training program. *Soc Int Urol J* 2022;**3**:62-8.

165. NHS England. *2019/20 National Cost Collection Data Publication*. NHS England; 2021. Available from: https://www.england.nhs.uk/publication/2019-20-national-cost-collection-data-publication/ [accessed 27th September 2022].

166. NHS England. *2020/2021 National Cost Collection data* NHS England; 2022. Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#ncc1819 [accessed 27th September 2022].

167. Imperial College London. *Imperial Prostate 7 - prostate assessment using comparative interventions - fast MRI and image-fusion for cancer (IP7-PACIFIC)*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2022. [accessed 30th October 2022].

168. Kruschke JK. *Doing Bayesian data analysis: a tutorial with R and BUGS*. London: Academic Press; 2010.

169. Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Med Decis Making* 2003;**23**:341-50.

170. Cornud F, Lefevre A, Flam T, Dumonceau O, Galiano M, Soyer P, et al. MRI-directed high-frequency (29MhZ) TRUS-guided biopsies: initial results of a single-center study. *Eur Radiol* 2020;**30**:4838-46.

171. Klotz L, Lughezzani G, Maffei D, Sanchez A, Pereira JG, Staerman F, et al. Comparison of micro-ultrasound and multiparametric magnetic resonance imaging for prostate cancer: A multicenter, prospective analysis. *Can Urol Assoc J* 2021;**15**:E11-E6.

172. Wiemer L, Hollenbach M, Heckmann R, Kittner B, Plage H, Reimann M, et al. Evolution of targeted prostate biopsy by adding micro-ultrasound to the magnetic resonance imaging pathway. *Eur Urol Focus* 2021;**7**:1292-9.

173. Hofbauer SL, Luger F, Harland N, Plage H, Reimann M, Hollenbach M, et al. A non-inferiority comparative analysis of micro-ultrasonography and MRI-targeted biopsy in men at risk of prostate cancer. *BJU Int* 2022;**129**:648-54.

174. Perlis N, Lawendy B, Barkin J. How I Do It - MRI-ultrasound fusion prostate biopsy using the Fusion MR and Fusion Bx systems. *Can J Urol* 2020;**27**:10185-91.

175. Miah S, Hosking-Jervis F, Connor MJ, Eldred-Evans D, Shah TT, Arya M, et al. A Multicentre Analysis of the Detection of Clinically Significant Prostate Cancer Following Transperineal Image-fusion Targeted and Nontargeted Systematic Prostate Biopsy in Men at Risk. *Eur Urol Oncol* 2020;**3**:262-9.

176. Immerzeel J, Israel B, Bomers J, Schoots IG, van Basten JP, Kurth KH, et al. Multiparametric Magnetic Resonance Imaging for the Detection of Clinically Significant Prostate Cancer: What Urologists Need to Know. Part 4: Transperineal Magnetic Resonance-Ultrasound Fusion Guided Biopsy Using Local Anesthesia. *Eur Urol* 2022;**81**:110-7.

177. Oh J, Tyldesley S, Pai HH, McKenzie MR, Halperin RM, Duncan GG, et al. An updated analysis of survival endpoints for ASCENDE-RT, a randomized trial comparing a low-dose-rate brachytherapy boost to a dose-escalated external beam boost for high- and intermediate-risk prostate cancer. *Int J Radiat Oncol Biol Phys* 2020;**108**:S62.

178. Hedley M. Multi-parametric magnetic resonance imaging before prostate biopsy. *Cancer Nursing Practice* 2017;**16**:21-5.

179. de Vries KC, Wortel RC, Oomen-de Hoop E, Heemsbergen WD, Pos FJ, Incrocci L. Hyprofractionated versus conventionally fractionated radiation therapy for patients with intermediate- or high-risk, localized, prostate cancer: 7-year outcomes from the randomized, multicenter, open-label, phase 3 HYPRO trial. *Int J Radiat Oncol Biol Phys* 2020;**106**:108-15.

180. David PD, Clare G, Isabel S, Vincent K, Alison Jane B, Ananya C, et al. Eight-year outcomes of a phase III randomized trial of conventional versus hypofractionated high-dose intensity modulated radiotherapy for prostate cancer (CRUK/06/016): update from the CHHiP Trial. *J Clin Oncol* 2020;**38**:325.

181. Fransson P, Nilsson P, Gunnlaugsson A, Beckman L, Tavelin B, Norman D, et al. Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer (HYPO-RT-PC): patient-reported quality-of-life outcomes of a randomised, controlled, non-inferiority, phase 3 trial. *Lancet Oncol* 2021;**22**:235-45.

182. Wilt TJ, Vo TN, Langsetmo L, Dahm P, Wheeler T, Aronson WJ, et al. Radical prostatectomy or observation for clinically localized prostate cancer: extended follow-up of the prostate cancer intervention versus observation trial (PIVOT). *Eur Urol* 2020;**77**:713-24.

183. Bill-Axelson A, Holmberg L, Garmo H, Taari K, Busch C, Nordling S, et al. Radical prostatectomy or watchful waiting in prostate cancer - 29-year follow-up. *N Engl J Med* 2018;**379**:2319-29.

184. Fizazi K, Carmel A, Joly F, Delva R, Gravis G, Rolland F, et al. Updated results of GETUG-12, a phase III trial of docetaxel-based chemotherapy in high-risk localized prostate cancer, with a 12-year follow-up. *Ann Oncol* 2018;**29**:viii271.

185. Pepe P, Aragona F. Morbidity after transperineal prostate biopsy in 3000 patients undergoing 12 vs 18 vs more than 24 needle cores. *Urology* 2013;**81**:1142-6.

186. Tamhankar AS, El-Taji O, Vasdev N, Foley C, Popert R, Adshead J. P9-7 The clinical and financial implications of a decade of prostate biopsies in the NHS - interrogation of the Hospital Episode Statistics (HES) Data 2008-2019 [BAUS abstracts]. *J Clin Urol* 2020;**13**:62.

187. NHS England. *2018/19 National Cost Collection Data Publication*. NHS England; 2021. Available from: https://www.england.nhs.uk/publication/2018-19-national-cost-collection-data-publication/ [accessed 27th September 2022].

188. *NPCA Organisational Audit 2021. Table of NHS provider-level results*. National Prostate Cancer Audit, The Royal College of Surgeons of England; 2022. Available from: https://www.npca.org.uk/reports/npca-organisational-audit-2021/ [accessed 29th September 2022].

189. National Institute for Health and Care Excellence. *NICE impact prostate cancer*. London: NICE; 2020.

190. *Drugs and pharmaceutical electronic market information tool (eMIT)*. Deparment of Health and Social Care; 2021. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed 29th September 2022].

191. *British National Formulary (BNF)*. BMJ Group and the Royal Pharmaceutical Society of Great Britain; 2022. Available from: https://bnf.nice.org.uk/ [accessed 27th September 2022].

APPENDICES

APPENDIX 1. LITERATURE SEARCH STRATEGIES

Database search strategies

MEDLINE ALL

(includes: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid http://ovidsp.ovid.com/

Date range: 1946 to May 13, 2022

Date searched: 16th May 2022

Records retrieved: 3129

MEDLINE ALL was searched again on 2nd August 2022. 3218 studies were retrieved.

1 exp Prostatic Neoplasms/ (142378)
2 Prostatic Intraepithelial Neoplasia/ (1399)
3 ((prostate$ or prostatic or intraprostatic) adj4 (cancer$ or neoplas$ or tumour$ or tumor$ or malignan$ or metasta$ or carcinoma$ or adenocarcinoma$ or lesion$ or nodul$ or sarcoma$ or lymphoma$)).ti,ab. (165600)
4 (PCa or sPCa or csPCa or PrCa).ti,ab. (52571)
5 (((atypical adj3 proliferation) or ASAP) and prostat$).mp. (292)
6 or/1-5 (224791)
7 Magnetic Resonance Imaging/ (453356)
8 Multiparametric Magnetic Resonance Imaging/ (961)
9 (magnetic resonance or MRI or MR imag$ or MR scan$).ti,ab. (560471)
10 (mpMRI or mp-MRI or mpMR imag$ or mpMR scan$ or mp-MR imag$ or mp-MR scan$ or bpMRI or bp-MRI or bpMR imag$ or bpMR scan$ or bp-MR imag$ or bp-MR scan$).ti,ab. (2060)
11 or/7-10 (721668)
12 Image Interpretation, Computer-Assisted/ (47627)
13 (fusion$ or fuse$ or fusing$).ti,ab. (299284)
14 cognitive$.ti,ab. (424900)
15 (visual$ adj3 (estimat$ or direct$ or align$ or guid$ or influenc$)).ti,ab. (28436)
16 registration$.ti,ab. (161125)
17 (elastic or rigid or nonrigid).ti,ab. (138219)
18 Software/ (120348)
19 (software or hardware).ti,ab. (224399)
20 or/12-19 (1355053)
21 Prostate/ (39209)
22 (prostate$ or prostatic).ti,ab. (234214)
23 21 or 22 (238231)
24 Biopsy/ (185156)
25 Image-Guided Biopsy/ (5020)
26 Endoscopic Ultrasound-Guided Fine Needle Aspiration/ (3254)
27 Biopsy, Fine-Needle/ (14970)
28 Biopsy, Large-Core Needle/ (2307)
29 Biopsy, Needle/ (49647)
30 (biopsy or biopsie$ or rebiopsy or rebiopsie$).ti,ab. (427177)
31 or/24-30 (548867)
32 23 and 31 (26179)
33 6 and 11 and 20 and 32 (1621)
34 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (biopsy or biopsie$)).ti,ab. (860)
35 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 prebiops$).ti,ab. (160)
36 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (ultrasound$ or ultrasonic$ or ultrasonograph$ or TRUS or transperineal$ or transrectal$)).ti,ab. (773)
37 or/34-36 (1626)
38 6 and 37 (662)
39 (target$ adj4 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).ti,ab. (3800)
40 (focal adj2 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).ti,ab. (636)
41 39 or 40 (4405)
42 6 and 41 (1842)
43 (target$ adj4 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).ti,ab. (4003)
44 (focal adj2 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).ti,ab. (546)
45 43 or 44 (4534)
46 6 and 32 and 45 (1125)

47 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj3 (guid$ or influenc$ or direct$ or align$)).ti,ab. (11942)
48 6 and 32 and 47 (951)
49 ((MRI stratified or magnetic resonance imaging stratified) adj3 pathway$).ti,ab. (3)
50 33 or 38 or 42 or 46 or 48 or 49 (3265)
51 (MRGB or MR-GB or MRIGB or MRI-GB).ti,ab. (75)
52 (MRIFB or MRI-FB).ti,ab. (3)
53 (MRFTB or MRF-TB).ti,ab. (9)
54 (MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx).ti,ab. (96)
55 FBx.ti,ab. (94)
56 (FUSTB or FUS-TB or TB-FUS).ti,ab. (9)
57 Fusion TB.ti,ab. (21)
58 (MRI-TRUS or MRI-TRUSB or MRI-TPB).ti,ab. (189)
59 (COG-TB or TB-COG or CBx).ti,ab. (530)
60 TRUS-TB.ti,ab. (3)
61 ("MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB").ti,ab. (306)
62 or/51-61 (1105)
63 6 and 62 (437)
64 50 or 63 (3292)
65 (fusion$ adj3 (software or hardware or computer$ or device$ or system$ or technolog$ or machine$ or platform$)).ti,ab. (5680)
66 6 and 65 (294)
67 64 or 66 (3331)
68 KOELIS.ti,ab. (23)
69 Fusion Bx.ti,ab. (1)
70 Biojet.ti,ab. (28)
71 (Trinity or PROMAP).ti,ab. (1329)
72 Fusion MR.ti,ab. (8)
73 (bkFusion or bk Fusion or BK3000 or BK 3000 or BK5000 or BK 5000 or Predictive Fusion).ti,ab. (7)
74 or/70-73 (1371)
75 6 and 74 (20)
76 68 or 69 or 75 (38)
77 Biopsee.ti,ab. (6)
78 UroNav.ti,ab. (17)
79 ("iSR'obot" or iSRobot or iSR obot or UroFusion or UroBiopsy).ti,ab. (2)
80 (FusionVu$ or ExactVu$).ti,ab. (12)
81 DynaCAD.ti,ab. (9)
82 (ARTEMIS or ProFuse).ti,ab. (4760)
83 Mona Lisa.ti,ab. (106)
84 or/81-83 (4874)
85 6 and 84 (54)
86 or/77-80 (34)
87 85 or 86 (81)
88 67 or 76 or 87 (3362)
89 exp animals/ not humans.sh. (5007245)
90 88 not 89 (3357)
91 limit 90 to yr="2008 -Current" (3129)

Key:

/ = subject heading (MeSH heading)

sh = subject heading (MeSH heading)

exp = exploded subject heading (MeSH heading)

$ = truncation

ti,ab = terms in title or abstract fields

mp = multi-purpose field search – terms in title, original title, abstract, name of substance word, or subject heading word

adj3 = terms within three words of each other (any order)

Cochrane Controlled Register of Trials (CENTRAL)

via Wiley http://onlinelibrary.wiley.com/

Issue: Issue 4 of 12, April 2022

Date searched: 16th May 2022

Records retrieved: 425

CENTRAL was searched again on 2nd August 2022. 434 studies were retrieved.

#1 MeSH descriptor: [Prostatic Neoplasms] explode all trees 6115
#2 MeSH descriptor: [Prostatic Intraepithelial Neoplasia] this term only 47
#3 ((prostate* or prostatic or intraprostatic) near/4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*)):ti,ab,kw 15719
#4 (PCa or sPCa or csPCa or PrCa):ti,ab,kw 5554
#5 (((atypical near/3 proliferation) or ASAP) and prostat*):ti,ab,kw 21
#6 #1 or #2 or #3 or #4 or #5 20099
#7 MeSH descriptor: [Magnetic Resonance Imaging] this term only 7831
#8 MeSH descriptor: [Multiparametric Magnetic Resonance Imaging] this term only 11
#9 ("magnetic resonance" or MRI or (MR next imag*) or (MR next scan*)):ti,ab,kw 41256
#10 (mpMRI or mp-MRI or (mpMR next imag*) or (mpMR next scan*) or (mp-MR next imag*) or mp-MR scan* or bpMRI or bp-MRI or (bpMR next imag*) or (bpMR next scan*) or bp-MR imag* or bp-MR scan*):ti,ab,kw 260
#11 #7 or #8 or #9 or #10 41264
#12 MeSH descriptor: [Image Interpretation, Computer-Assisted] this term only 875
#13 (fusion* or fuse* or fusing*):ti,ab,kw 8635
#14 cognitive*:ti,ab,kw 80126
#15 (visual* near/3 (estimat* or direct* or align* or guid* or influenc*)):ti,ab,kw 2089
#16 registration*:ti,ab,kw 66768
#17 (elastic or rigid or nonrigid):ti,ab,kw 6102
#18 MeSH descriptor: [Software] this term only 1008
#19 (software or hardware):ti,ab,kw 26282
#20 #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 180581
#21 MeSH descriptor: [Prostate] this term only 975
#22 (prostate* or prostatic):ti,ab,kw 23298
#23 #21 or #22 23298
#24 MeSH descriptor: [Biopsy] this term only 3365
#25 MeSH descriptor: [Image-Guided Biopsy] this term only 119
#26 MeSH descriptor: [Endoscopic Ultrasound-Guided Fine Needle Aspiration] this term only 156
#27 MeSH descriptor: [Biopsy, Needle] explode all trees 1270
#28 (biopsy or biopsie* or rebiopsy or rebiopsie*):ti,ab,kw 32970
#29 #24 or #25 or #26 or #27 or #28 33007

#30 #23 and #29 2832
#31 #6 and #11 and #20 and #30 211
#32 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) near/6 (prior or previous* or preced* or before* or earlier or first or initial*) near/6 (biopsy or biopsie*)):ti,ab,kw 95
#33 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) near/6 prebiops*):ti,ab,kw 34
#34 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) near/6 (prior or previous* or preced* or before* or earlier or first or initial*) near/6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*)):ti,ab,kw 84
#35 (target* near/4 (biopsy or biopsie* or rebiopsy or rebiopsie*)):ti,ab,kw 573
#36 (focal near/2 (biopsy or biopsie* or rebiopsy or rebiopsie*)):ti,ab,kw 22
#37 #32 or #33 or #34 or #35 or #36 715
#38 #6 and #37 324
#39 (target* near/4 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI)):ti,ab,kw 453
#40 (focal near/2 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI)):ti,ab,kw 38
#41 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) near/3 (guid* or influenc* or direct* or align*)):ti,ab,kw 923
#42 #39 or #40 or #41 1299
#43 #6 and #30 and #42 279
#44 (("MRI stratified" or "magnetic resonance imaging stratified") near/3 pathway*):ti,ab,kw 0
#45 #31 or #38 or #43 or #44 430
#46 (MRGB or MR-GB or MRIGB or MRI-GB or MRIFB or MRI-FB or MRFTB or MRF-TB or MRFTB or MRF-TB or MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx or FBx or FUSTB or FUS-TB or TB-FUS or "Fusion TB" or MRI-TRUS or MRI-TRUSB or MRI-TPB or COG-TB or TB-COG or CBx or TRUS-TB or "MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB"):ti,ab,kw 136
#47 #6 and #46 82
#48 #45 or #47 431
#49 (fusion* near/3 (software or hardware or computer* or device* or system* or technolog* or machine* or platform*)):ti,ab,kw 267
#50 #6 and #49 57
#51 #48 or #50 434
#52 (KOELIS or "Fusion Bx" or Biojet):ti,ab,kw 18
#53 (Trinity or PROMAP or "Fusion MR" or bkFusion or "bk Fusion" or BK3000 or "BK 3000" or BK5000 or "BK 5000" or "Predictive Fusion"):ti,ab,kw 161
#54 #6 and #53 3
#55 #52 or #54 19
#56 (Biopsee or UroNav or "iSR'obot" or iSRobot or "iSR obot" or UroFusion or UroBiopsy or FusionVu* or ExactVu*):ti,ab,kw 19
#57 (DynaCAD or ARTEMIS or ProFuse or "Mona Lisa"):ti,ab,kw 283
#58 #6 and #57 9
#59 #56 or #58 27
#60 #51 or #55 or #59 with Publication Year from 2008 to 2022, in Trials 425
#61 #51 or #55 or #59 in Cochrane Reviews, Cochrane Protocols 1

Key:

MeSH descriptor = subject heading (MeSH heading)

* = truncation

ti,ab,kw = terms in title, abstract or keyword fields

near/3 = terms within three words of each other (any order)

next = terms are next to each other

Cochrane Database of Systematic Reviews (CDSR)

via Wiley http://onlinelibrary.wiley.com/

Issue: Issue 5 of 12, May 2022

Date searched: 16th May 2022

Records retrieved: 1

See above under CENTRAL for search strategy.

Cumulative Index to Nursing & Allied Health (CINAHL Plus)

via Ebsco http://onlinelibrary.wiley.com/

Date range: Inception to 20220516

Date searched: 16th May 2022

Records retrieved: 916

S1 (MH "Prostatic Neoplasms+") 34,206
S2 TI ((prostate* or prostatic or intraprostatic) N4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*)) OR AB ((prostate* or prostatic or intraprostatic) N4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*)) 35,654
S3 TI ((PCa or sPCa or csPCa or PrCa)) OR AB ((PCa or sPCa or csPCa or PrCa)) 7,320
S4 TI (((atypical N3 proliferation) or ASAP) and prostat*) OR AB (((atypical N3 proliferation) or ASAP) and prostat*)29
S5 S1 OR S2 OR S3 OR S4 48,489
S6 (MH "Magnetic Resonance Imaging") 136,332
S7 TI (("magnetic resonance" or MRI or (MR N1 imag*) or (MR N1 scan*))) OR AB (("magnetic resonance" or MRI or (MR N1 imag*) or (MR N1 scan*))) 123,908
S8 TI (((mpMRI or mp-MRI or (mpMR N1 imag*) or (mpMR N1 scan*) or (mp-MR N1 imag*) or (mp-MR N1 scan*) or bpMRI or bp-MRI or (bpMR N1 imag*) or (bpMR N1 scan*) or (bp-MR N1 imag*) or (bp-MR N1 scan*))) OR AB (((mpMRI or mp-MRI or (mpMR N1 imag*) or (mpMR N1 scan*) or (mp-MR N1 imag*) or (mp-MR N1 scan*) or bpMRI or bp-MRI or (bpMR N1 imag*) or (bpMR N1 scan*) or (bp-MR N1 imag*) or (bp-MR N1 scan*))) 631
S9 S6 OR S7 OR S8 181,020
S10 (MH "Image Interpretation, Computer Assisted") 9,454
S11 TI (fusion* or fuse* or fusing*) OR AB (fusion* or fuse* or fusing*) 26,160
S12 TI cognitive* OR AB cognitive* 154,740
S13 TI (visual* N3 (estimat* or direct* or align* or guid* or influenc*)) OR AB (visual* N3 (estimat* or direct* or align* or guid* or influenc*)) 4,578
S14 TI registration* OR AB registration* 64,987
S15 TI (elastic or rigid or nonrigid) OR AB (elastic or rigid or nonrigid) 12,473
S16 (MH "Software") 31,273
S17 TI (software or hardware) OR AB (software or hardware) 59,300
S18 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 341,260

S19 (MH "Prostate") 3,816
S20 TI (prostate* or prostatic) OR AB (prostate* or prostatic) 45,719
S21 S19 OR S20 46,101
S22 (MH "Biopsy") 35,975
S23 (MH "Biopsy, Needle") 11,989
S24 TI (biopsy or biopsie* or rebiopsy or rebiopsie*) OR AB (biopsy or biopsie* or rebiopsy or rebiopsie*) 59,743
S25 S22 OR S23 OR S24 84,744
S26 S21 AND S25 4,603
S27 S5 AND S9 AND S18 AND S26 463
S28 TI ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N6 (prior or previous* or preced* or before* or earlier or first or initial*) N6 (biopsy or biopsie*)) OR AB ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N6 (prior or previous* or preced* or before* or earlier or first or initial*) N6 (biopsy or biopsie*)) 254
S29 TI ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N6 prebiops*) OR AB ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N6 prebiops*) 45
S30 TI ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N6 (prior or previous* or preced* or before* or earlier or first or initial*) N6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*)) OR AB ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N6 (prior or previous* or preced* or before* or earlier or first or initial*) N6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*)) 289
S31 TI (target* N4 (biopsy or biopsie* or rebiopsy or rebiopsie*)) OR AB (target* N4 (biopsy or biopsie* or rebiopsy or rebiopsie*)) 961
S32 TI (focal N2 (biopsy or biopsie* or rebiopsy or rebiopsie*)) OR AB (focal N2 (biopsy or biopsie* or rebiopsy or rebiopsie*)) 136
S33 S28 OR S29 OR S30 OR S31 OR S32 1,512
S34 S5 AND S33 591
S35 TI (target* N4 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)) OR AB (target* N4 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)) 880
S36 TI (focal N2 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)) OR AB (focal N2 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)) 257
S37 TI ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N3 (guid* or influenc* or direct* or align*)) OR AB ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) N3 (guid* or influenc* or direct* or align*)) 3,420
S38 S35 OR S36 OR S37 4,296
S39 S5 AND S26 AND S38 533
S40 TI (("MRI stratified" or "magnetic resonance imaging stratified") N3 pathway*) OR AB (("MRI stratified" or "magnetic resonance imaging stratified") N3 pathway*) 2
S41 S27 OR S34 OR S39 OR S40 909
S42 TI (MRGB or MR-GB or MRIGB or MRI-GB or MRIFB or MRI-FB or MRFTB or MRF-TB or MRFTB or MRF-TB or MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx or FBx or FUSTB or FUS-TB or TB-FUS or "Fusion TB" or MRI-TRUS or MRI-TRUSB or MRI-TPB or COG-TB or TB-COG or CBx or TRUS-TB or "MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB") OR AB (MRGB or MR-GB or MRIGB or MRI-GB or MRIFB or MRI-FB or MRFTB or MRF-TB or MRFTB or MRF-TB or MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx or FBx or FUSTB or FUS-TB or TB-FUS or "Fusion TB" or MRI-TRUS or MRI-TRUSB or MRI-TPB or COG-TB or TB-COG or CBx or TRUS-TB or "MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB") 185
S43 S5 AND S42 126
S44 S41 OR S43 915

S45 TI (fusion* N3 (software or hardware or computer* or device* or system* or technolog* or machine* or platform*)) OR AB (fusion* N3 (software or hardware or computer* or device* or system* or technolog* or machine* or platform*)) 832
S46 S5 AND S45 86
S47 S44 OR S46 922
S48 TI (KOELIS or "Fusion Bx" or Biojet) OR AB (KOELIS or "Fusion Bx" or Biojet) 17
S49 TI (Trinity or PROMAP or "Fusion MR" or bkFusion or "bk Fusion" or BK3000 or "BK 3000" or BK5000 or "BK 5000" or "Predictive Fusion") OR AB (Trinity or PROMAP or "Fusion MR" or bkFusion or "bk Fusion" or BK3000 or "BK 3000" or BK5000 or "BK 5000" or "Predictive Fusion") 482
S50 S5 AND S49 2
S51 S48 OR S50 18
S52 TI (Biopsee or UroNav or "iSR'obot" or iSRobot or "iSR obot" or UroFusion or UroBiopsy or FusionVu* or ExactVu*) OR AB (Biopsee or UroNav or "iSR'obot" or iSRobot or "iSR obot" or UroFusion or UroBiopsy or FusionVu* or ExactVu*) 11
S53 TI (DynaCAD or ARTEMIS or ProFuse or "Mona Lisa") AND AB (DynaCAD or ARTEMIS or ProFuse or "Mona Lisa") 32
S54 S5 AND S53 0
S55 S52 OR S54 11
S56 S47 OR S51 OR S55 925
S57 S47 OR S51 OR S55 Limiters - Published Date: 20080101-20221231 916

Key:

MH = CINAHL subject heading

+ = exploded CINAHL subject heading

* = truncation

TI = terms in the title

AB = terms in the abstract

N3 = terms within three words of each other (any order)

Database of Abstracts of Reviews of Effects (DARE)

via http://onlinelibrary.wiley.com/

Date range: Inception – 31st March 2015

Date searched: 16th May 2022

Records retrieved: 7

1 MeSH DESCRIPTOR Prostatic Neoplasms EXPLODE ALL TREES 709
2 MeSH DESCRIPTOR Prostatic Intraepithelial Neoplasia 2
3 ((prostate* or prostatic or intraprostatic) NEAR4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*)) 891
4 (PCa or sPCa or csPCa or PrCa) 44
5 ((atypical NEAR3 proliferation) or ASAP) AND (prostat*) 1
6 #1 OR #2 OR #3 OR #4 OR #5 935
7 MeSH DESCRIPTOR Magnetic Resonance Imaging 693
8 MeSH DESCRIPTOR Multiparametric Magnetic Resonance Imaging 0
9 ("magnetic resonance" or MRI or MR imag* or MR scan*) 1337

10 (mpMRI or mp-MRI or mpMR imag* or mpMR scan* or mp-MR imag* or mp-MR scan* or bpMRI or bp-MRI or bpMR imag* or bpMR scan* or bp-MR imag* or bp-MR scan*) 2
11 #7 OR #8 OR #9 OR #10 1337
12 MeSH DESCRIPTOR Image Interpretation, Computer-Assisted 27
13 (fusion* or fuse* or fusing* or cognitive or registration* or elastic or rigid or nonrigid) 3376
14 (visual* NEAR3 (estimat* or direct* or align* or guid* or influenc*)) 23
15 MeSH DESCRIPTOR Software 76
16 (software or hardware) 812
17 #12 OR #13 OR #14 OR #15 OR #16 4163
18 MeSH DESCRIPTOR Prostate 82
19 (prostate* or prostatic) 1283
20 #18 OR #19 1283
21 MeSH DESCRIPTOR Biopsy 248
22 MeSH DESCRIPTOR Image-Guided Biopsy 11
23 MeSH DESCRIPTOR Endoscopic Ultrasound-Guided Fine Needle Aspiration 19
24 MeSH DESCRIPTOR Biopsy, Fine-Needle 83
25 MeSH DESCRIPTOR Biopsy, Large-Core Needle 8
26 MeSH DESCRIPTOR Biopsy, Needle 164
27 (biopsy or biopsie* or rebiopsy or rebiopsie*) 1457
28 #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 1473
29 #20 AND #28 137
30 #6 AND #11 AND #17 AND #29 4
31 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR6 (biopsy or biopsie*)) 39
32 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR6 prebiops*) 0
33 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*)) 121
34 (target* NEAR4 (biopsy or biopsie* or rebiopsy or rebiopsie*)) 10
35 (focal NEAR2 (biopsy or biopsie* or rebiopsy or rebiopsie*)) 0
36 #31 OR #32 OR #33 OR #34 OR #35 155
37 #6 AND #36 10
38 (target* NEAR4 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI)) 3
39 (focal* NEAR2 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI)) 1
40 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR3 (guid* or influenc* or direct* or align*)) 65
41 #38 OR #39 OR #40 67
42 #6 AND #29 AND #41 5
43 (("MRI stratified" or "magnetic resonance imaging stratified") NEAR3 pathway*) 0
44 (MRGB or MR-GB or MRIGB or MRI-GB or MRIFB or MRI-FB or MRFTB or MRF-TB or MRFTB or MRF-TB or MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx or FBx or FUSTB or FUS-TB or TB-FUS or "Fusion TB" or MRI-TRUS or MRI-TRUSB or MRI-TPB or COG-TB or TB-COG or CBx or TRUS-TB or "MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB") 1
45 (fusion* NEAR3 (software or hardware or computer* or device* or system* or technolog* or machine* or platform*)) 34
46 #44 OR #45 35
47 #6 AND #46 2
48 #30 OR #37 OR #42 OR #43 OR #47 11
49 (Trinity or PROMAP or "Fusion MR" or bkFusion or "bk Fusion" or BK3000 or "BK 3000" or BK5000 or "BK 5000" or "Predictive Fusion") 2
50 (DynaCAD or ARTEMIS or ProFuse or "Mona Lisa") 8

51 #49 OR #50 10
52 #6 AND #51 0
53 (KOELIS or "Fusion Bx" or Biojet or Biopsee or UroNav or "iSR'obot" or iSRobot or "iSR obot" or UroFusion or UroBiopsy or FusionVu* or ExactVu*) 0
54 #48 OR #52 OR #53 11

Key:

MeSH DESCRIPTOR = subject heading (MeSH heading)

* = truncation

NEAR3 = terms within three words of each other (order specified)

EconLit

via Ovid http://ovidsp.ovid.com/

Date range: 1886 to May 05, 2022

Date searched: 16th May 2022

Records retrieved: 0

1 ((prostate$ or prostatic or intraprostatic) adj4 (cancer$ or neoplas$ or tumour$ or tumor$ or malignan$ or metasta$ or carcinoma$ or adenocarcinoma$ or lesion$ or nodul$ or sarcoma$ or lymphoma$)).mp. (114)
2 (PCa or sPCa or csPCa or PrCa).mp. (541)
3 (((atypical adj3 proliferation) or ASAP) and prostat$).mp. (0)
4 or/1-3 (651)
5 (magnetic resonance or MRI or MR imag$ or MR scan$).mp. (188)
6 (mpMRI or mp-MRI or mpMR imag$ or mpMR scan$ or mp-MR imag$ or mp-MR scan$ or bpMRI or bp-MRI or bpMR imag$ or bpMR scan$ or bp-MR imag$ or bp-MR scan$).mp. (0)
7 5 or 6 (188)
8 (fusion$ or fuse$ or fusing$).mp. (643)
9 cognitive$.mp. (17030)
10 (visual$ adj3 (estimat$ or direct$ or align$ or guid$ or influenc$)).mp. (75)
11 registration$.mp. (1925)
12 (elastic or rigid or nonrigid).mp. (4352)
13 (software or hardware).mp. (15832)
14 or/8-13 (39541)
15 (prostate$ or prostatic).mp. (141)
16 (biopsy or biopsie$ or rebiopsy or rebiopsie$).mp. (17)
17 15 and 16 (4)
18 4 and 7 and 14 and 17 (0)
19 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (biopsy or biopsie$)).mp. (0)
20 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 prebiops$).mp. (0)
21 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (ultrasound$ or ultrasonic$ or ultrasonograph$ or TRUS or transperineal$ or transrectal$)).mp. (0)
22 (target$ adj4 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).mp. (0)
23 (focal adj2 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).mp. (0)
24 (target$ adj4 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).mp. (2)

25 4 and 24 (0)
26 (focal adj2 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).mp. (0)
27 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj3 (guid$ or influenc$ or direct$ or align$)).mp. (9)
28 4 and 27 (0)
29 ((MRI stratified or magnetic resonance imaging stratified) adj3 pathway$).mp. (0)
30 (MRGB or MR-GB or MRIGB or MRI-GB).mp. (0)
31 (MRIFB or MRI-FB).mp. (0)
32 (MRFTB or MRF-TB).mp. (0)
33 (MRFTB or MRF-TB).mp. (0)
34 (MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx).mp. (0)
35 FBx.mp. (0)
36 (FUSTB or FUS-TB or TB-FUS).mp. (0)
37 Fusion TB.mp. (0)
38 (MRI-TRUS or MRI-TRUSB or MRI-TPB).mp. (0)
39 (COG-TB or TB-COG or CBx).mp. (1)
40 4 and 39 (0)
41 TRUS-TB.mp. (0)
42 ("MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB").mp. (0)
43 (fusion$ adj3 (software or hardware or computer$ or device$ or system$ or technolog$ or machine$ or platform$)).mp. (26)
44 4 and 43 (0)
45 (KOELIS or Fusion Bx).mp. (0)
46 (Biojet or Trinity or PROMAP or Fusion MR or bkFusion or bk Fusion or BK3000 or BK 3000 or BK5000 or BK 5000 or Predictive Fusion).mp. (356)
47 4 and 46 (0)
48 (Biopsee or UroNav or "iSR'obot" or iSRobot or iSR obot or UroFusion or UroBiopsy or FusionVu$ or ExactVu$).mp. (0)
49 (DynaCAD or ARTEMIS or ProFuse or Mona Lisa).mp. (24)
50 4 and 49 (0)
51 18 or 19 or 20 or 21 or 22 or 23 or 25 or 26 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 40 or 41 or 42 or 44 or 45 or 47 or 48 or 50 (0)

Key:

$ = truncation

mp = multi-purpose field search – terms in title, original title, abstract, name of substance word, or subject heading word

adj3 = terms within three words of each other (any order)

Embase

via Ovid http://onlinelibrary.wiley.com/

Date range: 1974 to 2022 May 13

Date searched: 16th May 2022

Records retrieved: 6221

Embase was searched again on 2nd August 2022. After conference abstracts were removed, 3318 studies were retrieved.

1 exp prostate tumor/ (271321)
2 ((prostate$ or prostatic or intraprostatic) adj4 (cancer$ or neoplas$ or tumour$ or tumor$ or malignan$ or metasta$ or carcinoma$ or adenocarcinoma$ or lesion$ or nodul$ or sarcoma$ or lymphoma$)).ti,ab. (244110)
3 (PCa or sPCa or csPCa or PrCa).ti,ab. (77312)
4 (((atypical adj3 proliferation) or ASAP) and prostat$).mp. (644)
5 or/1-4 (351675)
6 nuclear magnetic resonance imaging/ (903701)
7 multiparametric magnetic resonance imaging/ (6477)
8 (magnetic resonance or MRI or MR imag$ or MR scan$).ti,ab. (819806)
9 (mpMRI or mp-MRI or mpMR imag$ or mpMR scan$ or mp-MR imag$ or mp-MR scan$ or bpMRI or bp-MRI or bpMR imag$ or bpMR scan$ or bp-MR imag$ or bp-MR scan$).ti,ab. (4373)
10 or/6-9 (1163227)
11 computer assisted diagnosis/ (41296)
12 (fusion$ or fuse$ or fusing$).ti,ab. (361890)
13 cognitive$.ti,ab. (585952)
14 (visual$ adj3 (estimat$ or direct$ or align$ or guid$ or influenc$)).ti,ab. (35918)
15 registration$.ti,ab. (163670)
16 (elastic or rigid or nonrigid).ti,ab. (152621)
17 software/ or imaging software/ or nuclear magnetic resonance scanner software/ or ultrasound imaging system software/ (139562)
18 (software or hardware).ti,ab. (363110)
19 or/11-18 (1695591)
20 exp prostate/ (54557)
21 (prostate$ or prostatic).ti,ab. (336120)
22 20 or 21 (339264)
23 biopsy/ (174400)
24 image guided biopsy/ (6935)
25 endoscopic ultrasound guided fine needle biopsy/ (5968)
26 exp needle biopsy/ (79356)
27 biopsy technique/ (7739)
28 tumor biopsy/ (43525)
29 (biopsy or biopsie$ or rebiopsy or rebiopsie$).ti,ab. (685526)
30 or/23-29 (782015)
31 22 and 30 (43352)
32 prostate biopsy/ or exp transperineal biopsy/ or exp transrectal biopsy/ (24654)
33 31 or 32 (48987)
34 5 and 10 and 19 and 33 (3137)
35 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (biopsy or biopsie$)).ti,ab. (1707)
36 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 prebiops$).ti,ab. (248)
37 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (ultrasound$ or ultrasonic$ or ultrasonograph$ or TRUS or transperineal$ or transrectal$)).ti,ab. (1370)
38 or/35-37 (2954)
39 5 and 38 (1359)
40 (target$ adj4 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).ti,ab. (7633)
41 (focal adj2 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).ti,ab. (1195)
42 40 or 41 (8750)
43 5 and 40 (3525)
44 (target$ adj4 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).ti,ab. (6907)

45 (focal adj2 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).ti,ab. (959)
46 44 or 45 (7838)
47 5 and 31 and 46 (2297)
48 mri guided biopsy/ (246)
49 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj3 (guid$ or influenc$ or direct$ or align$)).ti,ab. (18601)
50 48 or 49 (18743)
51 5 and 31 and 50 (1937)
52 ((MRI stratified or magnetic resonance imaging stratified) adj3 pathway$).ti,ab. (3)
53 magnetic resonance imaging ultrasound fusion biopsy/ (128)
54 image guided noninferiority targeted biopsy/ (1)
55 cognitive biopsy/ (4)
56 software based targeted biopsy/ (1)
57 visually directed targeted biopsy/ (1)
58 ultrasound fusion targeted biopsy/ (3)
59 or/52-58 (140)
60 34 or 39 or 43 or 47 or 51 or 59 (6166)
61 (MRGB or MR-GB or MRIGB or MRI-GB).ti,ab. (132)
62 (MRIFB or MRI-FB).ti,ab. (8)
63 (MRFTB or MRF-TB).ti,ab. (36)
64 (MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx).ti,ab. (168)
65 FBx.ti,ab. (226)
66 (FUSTB or FUS-TB or TB-FUS).ti,ab. (11)
67 Fusion TB.ti,ab. (29)
68 (MRI-TRUS or MRI-TRUSB or MRI-TPB).ti,ab. (485)
69 (COG-TB or TB-COG or CBx).ti,ab. (829)
70 TRUS-TB.ti,ab. (8)
71 ("MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB").ti,ab. (777)
72 or/61-71 (2124)
73 5 and 72 (1009)
74 60 or 73 (6215)
75 (fusion$ adj3 (software or hardware or computer$ or device$ or system$ or technolog$ or machine$ or platform$)).ti,ab. (7446)
76 5 and 75 (707)
77 magnetic resonance imaging-ultrasound fusion-guided prostate biopsy device/ (245)
78 74 or 76 or 77 (6346)
79 KOELIS.ti,ab,dv. (180)
80 Fusion Bx.ti,ab,dv. (16)
81 Biojet.ti,ab,dv. (105)
82 (Trinity or PROMAP).ti,ab,dv. (2121)
83 Fusion MR.ti,ab,dv. (13)
84 (bkFusion or bk Fusion or BK3000 or BK 3000 or BK5000 or BK 5000 or Predictive Fusion).ti,ab,dv. (60)
85 or/81-84 (2295)
86 5 and 85 (148)
87 79 or 80 or 86 (307)
88 Biopsee.ti,ab,dv. (52)
89 UroNav.ti,ab,dv. (163)
90 ("iSR'obot" or iSRobot or iSR obot or UroFusion or UroBiopsy).ti,ab,dv. (31)
91 (FusionVu$ or ExactVu$).ti,ab,dv. (84)
92 DynaCAD.ti,ab,dv. (73)
93 (ARTEMIS or ProFuse).ti,ab,dv. (6586)
94 Mona Lisa.ti,ab,dv. (162)

95 or/92-94 (6817)
96 5 and 95 (247)
97 88 or 89 or 90 or 91 or 96 (506)
98 78 or 87 or 97 (6483)
99 (animal/ or animal experiment/ or animal model/ or animal tissue/ or nonhuman/) not exp human/ (6457016)
100 98 not 99 (6455)
101 limit 100 to yr="2008 -Current" (6221)

Key:

/ = subject heading (Emtree heading)

exp = exploded subject heading (Emtree heading)

$ = truncation

ti,ab = terms in title or abstract fields

mp = multi-purpose field search – terms in title, original title, abstract, name of substance word, or subject heading word

dv = terms in the device trade name field

adj3 = terms within three words of each other (any order)

Health Management and Information Consortium (HMIC)

via Ovid http://onlinelibrary.wiley.com/

Date range: 1979 to March 2022

Date searched: 16th May 2022

Records retrieved: 0

1 ((prostate$ or prostatic or intraprostatic) adj4 (cancer$ or neoplas$ or tumour$ or tumor$ or malignan$ or metasta$ or carcinoma$ or adenocarcinoma$ or lesion$ or nodul$ or sarcoma$ or lymphoma$)).mp. (736)
2 (PCa or sPCa or csPCa or PrCa).mp. (74)
3 (((atypical adj3 proliferation) or ASAP) and prostat$).mp. (0)
4 or/1-3 (792)
5 (magnetic resonance or MRI or MR imag$ or MR scan$).mp. (483)
6 (mpMRI or mp-MRI or mpMR imag$ or mpMR scan$ or mp-MR imag$ or mp-MR scan$ or bpMRI or bp-MRI or bpMR imag$ or bpMR scan$ or bp-MR imag$ or bp-MR scan$).mp. (0)
7 5 or 6 (483)
8 (fusion$ or fuse$ or fusing$).mp. (94)
9 cognitive$.mp. (2602)
10 (visual$ adj3 (estimat$ or direct$ or align$ or guid$ or influenc$)).mp. (23)
11 registration$.mp. (4038)
12 (elastic or rigid or nonrigid).mp. (258)
13 (software or hardware).mp. (1828)
14 or/8-13 (8757)
15 (prostate$ or prostatic).mp. (914)
16 (biopsy or biopsie$ or rebiopsy or rebiopsie$).mp. (303)
17 15 and 16 (36)
18 4 and 7 and 14 and 17 (0)

19 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (biopsy or biopsie$)).mp. (0)
20 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 prebiops$).mp. (0)
21 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous$ or preced$ or before$ or earlier or first or initial$) adj6 (ultrasound$ or ultrasonic$ or ultrasonograph$ or TRUS or transperineal$ or transrectal$)).mp. (0)
22 (target$ adj4 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).mp. (0)
23 (focal adj2 (biopsy or biopsie$ or rebiopsy or rebiopsie$)).mp. (0)
24 (target$ adj4 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).mp. (1)
25 4 and 24 (0)
26 (focal adj2 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI)).mp. (0)
27 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj3 (guid$ or influenc$ or direct$ or align$)).mp. (22)
28 4 and 27 (0)
29 ((MRI stratified or magnetic resonance imaging stratified) adj3 pathway$).mp. (0)
30 (MRGB or MR-GB or MRIGB or MRI-GB).mp. (0)
31 (MRIFB or MRI-FB).mp. (0)
32 (MRFTB or MRF-TB).mp. (0)
33 (MRFTB or MRF-TB).mp. (0)
34 (MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx).mp. (0)
35 FBx.mp. (0)
36 (FUSTB or FUS-TB or TB-FUS).mp. (0)
37 Fusion TB.mp. (0)
38 (MRI-TRUS or MRI-TRUSB or MRI-TPB).mp. (0)
39 (COG-TB or TB-COG or CBx).mp. (0)
40 TRUS-TB.mp. (0)
41 ("MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB").mp. (0)
42 (fusion$ adj3 (software or hardware or computer$ or device$ or system$ or technolog$ or machine$ or platform$)).mp. (1)
43 4 and 42 (0)
44 (KOELIS or Fusion Bx or Biojet).mp. (0)
45 (Trinity or PROMAP or Fusion MR or bkFusion or bk Fusion or BK3000 or BK 3000 or BK5000 or BK 5000 or Predictive Fusion).mp. (12)
46 4 and 45 (0)
47 (Biopsee or UroNav or "iSR'obot" or iSRobot or iSR obot or UroFusion or UroBiopsy or FusionVu$ or ExactVu$).mp. (0)
48 (DynaCAD or ARTEMIS or ProFuse or Mona Lisa).mp. (12)
49 4 and 48 (0)
50 18 or 19 or 20 or 21 or 22 or 23 or 25 or 26 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 43 or 44 or 46 or 47 or 49 (0)

Key:

$ = truncation

mp = multi-purpose field search – terms in title, original title, abstract, name of substance word, or subject heading word

adj3 = terms within three words of each other (any order)

Health Technology Assessment (HTA) database

via http://www.crd.york.ac.uk/CRDWeb/

Date range: Inception – 31st March 2018

Date searched: 16th May 2022

Records retrieved: 2

See under DARE for search strategy used.

International Health Technology Assessment (INAHTA) database

via http://onlinelibrary.wiley.com/

Date searched: 16th May 2022

Records retrieved: 38

1. ((((biopsy OR biopsie* OR rebiopsy OR rebiopsie*)[Title] OR (biopsy OR biopsie* OR rebiopsy OR rebiopsie*)[abs] OR (biopsy OR biopsie* OR rebiopsy OR rebiopsie*)[Keywords]) OR ("Biopsy, Needle"[mh]) OR ("Biopsy, Large-Core Needle"[mh]) OR ("Biopsy, Fine-Needle"[mh]) OR ("Endoscopic Ultrasound-Guided Fine Needle Aspiration"[mh]) OR ("Image-Guided Biopsy"[mh]) OR ("Biopsy"[mh])) AND (((prostate* OR prostatic)[Title] OR (prostate* OR prostatic)[abs] OR (prostate* OR prostatic)[Keywords]) OR ("Prostate"[mh]))) AND (((software OR hardware)[Title] OR (software OR hardware)[abs] OR (software OR hardware)[Keywords]) OR ("Software"[mh]) OR ((visual* AND (estimat* OR direct* OR align* OR guid* OR influenc*))[Title] OR (visual* AND (estimat* OR direct* OR align* OR guid* OR influenc*))[abs] OR (visual* AND (estimat* OR direct* OR align* OR guid* OR influenc*))[Keywords]) OR ((fusion* OR fuse* OR fusing* OR cognitive OR registration* OR elastic OR rigid OR nonrigid)[Title] OR (fusion* OR fuse* OR fusing* OR cognitive OR registration* OR elastic OR rigid OR nonrigid)[abs] OR (fusion* OR fuse* OR fusing* OR cognitive OR registration* OR elastic OR rigid OR nonrigid)[Keywords]) OR ("Image Interpretation, Computer-Assisted"[mh])) AND (((mpMRI OR mp-MRI OR mpMR imag* OR mpMR scan* OR mp-MR imag* OR mp-MR scan* OR bpMRI OR bp-MRI OR bpMR imag* OR bpMR scan* OR bp-MR imag* OR bp-MR scan*)[Title] OR (mpMRI OR mp-MRI OR mpMR imag* OR mpMR scan* OR mp-MR imag* OR mp-MR scan* OR bpMRI OR bp-MRI OR bpMR imag* OR bpMR scan* OR bp-MR imag* OR bp-MR scan*)[abs] OR (mpMRI OR mp-MRI OR mpMR imag* OR mpMR scan* OR mp-MR imag* OR mp-MR scan* OR bpMRI OR bp-MRI OR bpMR imag* OR bpMR scan* OR bp-MR imag* OR bp-MR scan*)[Keywords]) OR (("magnetic resonance" OR MRI OR MR imag* OR MR scan*)[Title] OR ("magnetic resonance" OR MRI OR MR imag* OR MR scan*)[abs] OR ("magnetic resonance" OR MRI OR MR imag* OR MR scan*)[Keywords]) OR ("Multiparametric Magnetic Resonance Imaging"[mh]) OR ("Magnetic Resonance Imaging"[mh])) AND (((ASAP AND prostat*)[Title] OR (ASAP AND prostat*)[abs] OR (ASAP AND prostat*)[Keywords]) OR ((atypical AND proliferation AND prostat*)[Title] OR (atypical AND proliferation AND prostat*)[abs] OR (atypical AND proliferation AND prostat*)[Keywords]) OR ((PCa OR sPCa OR csPCa OR PrCa)[Title] OR (PCa OR sPCa OR csPCa OR PrCa)[abs] OR (PCa OR sPCa OR csPCa OR PrCa)[Keywords]) OR (((cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Title] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[abs] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Keywords]) AND (((prostate* OR prostatic OR intraprostatic))[Title] OR ((prostate* OR prostatic OR intraprostatic))[abs] OR ((prostate* OR prostatic OR intraprostatic))[Keywords])) OR ("Prostatic Intraepithelial Neoplasia"[mh]) OR ("Prostatic Neoplasms"[mhe])) 4 hits

2. ((((target* OR focal) AND (biopsy OR biopsie* OR rebiopsy OR rebiopsie*))[Title] OR ((target* OR focal) AND (biopsy OR biopsie* OR rebiopsy OR rebiopsie*))[abs] OR ((target* OR focal) AND (biopsy OR biopsie* OR rebiopsy OR rebiopsie*))[Keywords]) OR (((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (ultrasound* OR ultrasonic* OR ultrasonograph* OR TRUS OR transperineal* OR transrectal*))[Title] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (ultrasound* OR ultrasonic* OR ultrasonograph* OR TRUS OR transperineal* OR transrectal*))[abs] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (ultrasound* OR ultrasonic* OR ultrasonograph* OR TRUS OR transperineal* OR transrectal*))[Keywords]) OR ((((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND prebiops*)[Title] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND prebiops*)[abs] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND prebiops*)[Keywords])) OR (((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND biops*)[Title] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND biops*)[abs] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND biops*)[Keywords])) AND (((ASAP AND prostat*)[Title] OR (ASAP AND prostat*)[abs] OR (ASAP AND prostat*)[Keywords]) OR ((atypical AND proliferation AND prostat*)[Title] OR (atypical AND proliferation AND prostat*)[abs] OR (atypical AND proliferation AND prostat*)[Keywords]) OR ((PCa OR sPCa OR csPCa OR PrCa)[Title] OR (PCa OR sPCa OR csPCa OR PrCa)[abs] OR (PCa OR sPCa OR csPCa OR PrCa)[Keywords]) OR (((cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Title] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[abs] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Keywords]) AND (((prostate* OR prostatic OR intraprostatic))[Title] OR ((prostate* OR prostatic OR intraprostatic))[abs] OR ((prostate* OR prostatic OR intraprostatic))[Keywords])) OR ("Prostatic Intraepithelial Neoplasia"[mh]) OR ("Prostatic Neoplasms"[mhe])) 9 hits

3. ((((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (guid* OR influenc* OR direct* OR align*))[Title] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (guid* OR influenc* OR direct* OR align*))[abs] OR ((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (guid* OR influenc* OR direct* OR align*))[Keywords]) OR (((target* OR focal) AND (MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI))[Title] OR ((target* OR focal) AND (MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI))[abs] OR ((target* OR focal) AND (MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI))[Keywords])) AND ((((biopsy OR biopsie* OR rebiopsy OR rebiopsie*)[Title] OR (biopsy OR biopsie* OR rebiopsy OR rebiopsie*)[abs] OR (biopsy OR biopsie* OR rebiopsy OR rebiopsie*)[Keywords]) OR ("Biopsy, Needle"[mh]) OR ("Biopsy, Large-Core Needle"[mh]) OR ("Biopsy, Fine-Needle"[mh]) OR ("Endoscopic Ultrasound-Guided Fine Needle Aspiration"[mh]) OR ("Image-Guided Biopsy"[mh]) OR ("Biopsy"[mh])) AND (((prostate* OR prostatic)[Title] OR (prostate* OR prostatic)[abs] OR (prostate* OR prostatic)[Keywords]) OR ("Prostate"[mh]))) AND (((ASAP AND prostat*)[Title] OR (ASAP AND prostat*)[abs] OR (ASAP AND prostat*)[Keywords]) OR ((atypical AND proliferation AND prostat*)[Title] OR (atypical AND proliferation AND prostat*)[abs] OR (atypical AND proliferation AND prostat*)[Keywords]) OR ((PCa OR sPCa OR csPCa OR PrCa)[Title] OR (PCa OR sPCa OR csPCa OR PrCa)[abs] OR (PCa OR sPCa OR csPCa OR PrCa)[Keywords]) OR (((cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Title] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[abs] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul*

OR sarcoma* OR lymphoma*)[Keywords]) AND (((prostate* OR prostatic OR intraprostatic))[Title] OR ((prostate* OR prostatic OR intraprostatic))[abs] OR ((prostate* OR prostatic OR intraprostatic))[Keywords])) OR ("Prostatic Intraepithelial Neoplasia"[mh]) OR ("Prostatic Neoplasms"[mhe])) 5 hits

4. ((MRGB OR MR-GB OR MRIGB OR MRI-GB OR MRIFB OR MRI-FB OR MRFTB OR MRF-TB OR MRFTB OR MRF-TB OR MRTB OR MR-TB OR MRITB OR MRI-TB OR MRTBx OR MR-TBx OR MRITBx OR MRI-TBx OR FBx OR FUSTB OR FUS-TB OR TB-FUS OR "Fusion TB" OR MRI-TRUS OR MRI-TRUSB OR MRI-TPB OR COG-TB OR TB-COG OR CBx OR TRUS-TB OR "MRI/TRUS" OR "mpMRI/TRUS" OR "MR/US" OR "MRI/TRUS-TB")[Title] OR (MRGB OR MR-GB OR MRIGB OR MRI-GB OR MRIFB OR MRI-FB OR MRFTB OR MRF-TB OR MRFTB OR MRF-TB OR MRTB OR MR-TB OR MRITB OR MRI-TB OR MRTBx OR MR-TBx OR MRITBx OR MRI-TBx OR FBx OR FUSTB OR FUS-TB OR TB-FUS OR "Fusion TB" OR MRI-TRUS OR MRI-TRUSB OR MRI-TPB OR COG-TB OR TB-COG OR CBx OR TRUS-TB OR "MRI/TRUS" OR "mpMRI/TRUS" OR "MR/US" OR "MRI/TRUS-TB")[abs] OR (MRGB OR MR-GB OR MRIGB OR MRI-GB OR MRIFB OR MRI-FB OR MRFTB OR MRF-TB OR MRFTB OR MRF-TB OR MRTB OR MR-TB OR MRITB OR MRI-TB OR MRTBx OR MR-TBx OR MRITBx OR MRI-TBx OR FBx OR FUSTB OR FUS-TB OR TB-FUS OR "Fusion TB" OR MRI-TRUS OR MRI-TRUSB OR MRI-TPB OR COG-TB OR TB-COG OR CBx OR TRUS-TB OR "MRI/TRUS" OR "mpMRI/TRUS" OR "MR/US" OR "MRI/TRUS-TB")[Keywords]) AND (((ASAP AND prostat*)[Title] OR (ASAP AND prostat*)[abs] OR (ASAP AND prostat*)[Keywords]) OR ((atypical AND proliferation AND prostat*)[Title] OR (atypical AND proliferation AND prostat*)[abs] OR (atypical AND proliferation AND prostat*)[Keywords]) OR ((PCa OR sPCa OR csPCa OR PrCa)[Title] OR (PCa OR sPCa OR csPCa OR PrCa)[abs] OR (PCa OR sPCa OR csPCa OR PrCa)[Keywords]) OR (((cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Title] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[abs] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Keywords]) AND (((prostate* OR prostatic OR intraprostatic))[Title] OR ((prostate* OR prostatic OR intraprostatic))[abs] OR ((prostate* OR prostatic OR intraprostatic))[Keywords])) OR ("Prostatic Intraepithelial Neoplasia"[mh]) OR ("Prostatic Neoplasms"[mhe])) 14 hits

5. ((Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa")[Title] OR (Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa")[abs] OR (Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa")[Keywords]) AND (((ASAP AND prostat*)[Title] OR (ASAP AND prostat*)[abs] OR (ASAP AND prostat*)[Keywords]) OR ((atypical AND proliferation AND prostat*)[Title] OR (atypical AND proliferation AND prostat*)[abs] OR (atypical AND proliferation AND prostat*)[Keywords]) OR ((PCa OR sPCa OR csPCa OR PrCa)[Title] OR (PCa OR sPCa OR csPCa OR PrCa)[abs] OR (PCa OR sPCa OR csPCa OR PrCa)[Keywords]) OR (((cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Title] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[abs] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Keywords]) AND (((prostate* OR prostatic OR intraprostatic))[Title] OR ((prostate* OR prostatic OR intraprostatic))[abs] OR ((prostate* OR prostatic OR intraprostatic))[Keywords])) OR ("Prostatic Intraepithelial Neoplasia"[mh]) OR ("Prostatic Neoplasms"[mhe]))

6. ((fusion* AND (software OR hardware OR computer* OR device* OR system* OR technolog* OR machine* OR platform*))[Title] OR (fusion* AND (software OR hardware OR computer* OR device* OR system* OR technolog* OR machine* OR platform*))[abs] OR (fusion* AND (software OR hardware OR computer* OR device* OR system* OR technolog* OR machine* OR platform*))[Keywords]) AND (((ASAP AND prostat*)[Title] OR (ASAP AND prostat*)[abs] OR (ASAP AND prostat*)[Keywords]) OR ((atypical AND proliferation AND prostat*)[Title] OR (atypical AND proliferation AND prostat*)[abs] OR (atypical AND proliferation AND prostat*)[Keywords]) OR ((PCa OR sPCa OR csPCa OR PrCa)[Title] OR (PCa OR sPCa OR csPCa OR PrCa)[abs] OR (PCa OR sPCa OR csPCa OR PrCa)[Keywords]) OR (((cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Title] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[abs] OR (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)[Keywords]) AND (((prostate* OR prostatic OR intraprostatic))[Title] OR ((prostate* OR prostatic OR intraprostatic))[abs] OR ((prostate* OR prostatic OR intraprostatic))[Keywords])) OR ("Prostatic Intraepithelial Neoplasia"[mh]) OR ("Prostatic Neoplasms"[mhe])) 2 hits

Key:

[abs] = abstract

[mh] = subject heading (MeSH heading)

[mhe] = exploded subject heading (MeSH heading)

* = truncation

Latin American and Caribbean Health Sciences Literature (LILACS)

via https://pesquisa.bvsalud.org/portal/advanced/?lang=en

Date searched: 16th May 2022

Records retrieved: 98

1. Search of title, abstract, subject heading fields: (prostat$) AND (cancer$ OR neoplas$ OR tumour$ OR tumor$ OR malignan$ OR metasta$ OR carcinoma$ OR adenocarcinoma$ OR lesion$ OR nodul$ OR sarcoma$ OR lymphoma$) AND ("magnetic resonance" OR MRI OR MR imag$ OR MR scan$ OR mpMRI OR mp-MRI OR mpMR imag$ OR mpMR scan$ OR mp-MR imag$ OR mp-MR scan$ OR bpMRI OR bp-MRI OR bpMR imag$ OR bpMR scan$ OR bp-MR imag$ OR bp-MR scan$) AND (fusion$ OR fuse$ OR fusing$ OR cognitive$ OR visual$ OR registration$ OR elastic OR rigid OR nonrigid OR software OR hardware OR target$ OR focal OR guid$ OR influenc$ OR direct$ OR align$) AND (biopsy OR biopsie$ OR rebiopsy OR rebiopsie$ OR prebiopsy$)
limit: 2008-2022
35 hits

2. Search of title, abstract, subject heading fields: (prostat$) AND (cancer$ OR neoplas$ OR tumour$ OR tumor$ OR malignan$ OR metasta$ OR carcinoma$ OR adenocarcinoma$ OR lesion$ OR nodul$ OR sarcoma$ OR lymphoma$) AND (MRI OR MR OR magnetic resonance OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) AND (ultrasound$ OR ultrasonic$ OR ultrasonograph$ OR TRUS OR transperineal$ OR transrectal$)
limit: 2008-2022
53 hits

3. Search of title, abstract, subject heading fields: (prostat$) AND (cancer$ OR neoplas$ OR tumour$ OR tumor$ OR malignan$ OR metasta$ OR carcinoma$ OR adenocarcinoma$ OR lesion$ OR nodul$ OR sarcoma$ OR lymphoma$) AND (MRGB OR MR-GB OR MRIGB OR MRI-GB OR MRIFB OR MRI-FB OR MRFTB OR MRF-TB OR MRFTB OR MRF-TB OR MRTB OR MR-TB OR MRITB OR MRI-TB OR MRTBx OR MR-TBx OR MRITBx OR MRI-TBx OR FBx OR FUSTB OR FUS-TB OR TB-FUS OR "Fusion TB" OR MRI-TRUS OR MRI-TRUSB OR MRI-TPB OR COG-TB OR TB-COG OR CBx OR TRUS-TB OR "MRI/TRUS" OR "mpMRI/TRUS" OR "MR/US" OR "MRI/TRUS-TB")
limit: 2008-2022
9 hits

4. Search of title, abstract, subject heading fields: (KOELIS OR "Fusion Bx" OR Biojet OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu$ OR ExactVu$)
limit: 2008-2022
0 hits

5. Search of title, abstract, subject heading fields: (Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa") AND (prostat$)
limit: 2008-2022
1 hit

NHS Economic Evaluations Database (NHS EED)

Via http://www.crd.york.ac.uk/CRDWeb/

Date range: Inception – 31st March 2015

Date searched: 16th May 2022

Records retrieved: 2

See under DARE for search strategy used.

Science Citation Index

via Web of Science, Clarivate Analytics https://clarivate.com/

Date range: 1900 - present

Date searched: 16th May 2022

Records retrieved: 3616

The Science Citation Index and the Conference Proceedings Citation Index-Science were both searched using the strategy below. Numbers of records retrieved are therefore the total number from searching both databases.

The Science Citation Index only was searched again on 2nd August 2022. 3561 studies were retrieved.

48 #45 OR #41 OR #37 3,616
47 #45 or #41 or #37 3,857

46 #45 OR #41 OR #37 3,857
45 #42 OR #44 69
44 #43 AND #4 42
43 TS=(DynaCAD or ARTEMIS or ProFuse or "Mona Lisa") 5,737
42 TS=(Biopsee or UroNav or "iSR'obot" or iSRobot or "iSR obot" or UroFusion or UroBiopsy or FusionVu* or ExactVu*) 34
41 #40 OR #38 41
40 #39 AND #4 19
39 TS=(Biojet or Trinity or PROMAP or "Fusion MR" or bkFusion or "bk Fusion" or BK3000 or "BK 3000" or BK5000 or "BK 5000" or "Predictive Fusion") 2,748
38 TS=(KOELIS or "Fusion Bx") 25
37 #36 OR #34 OR #32 3,825
36 #35 AND #4 471
35 TS=(fusion* NEAR/3 (software or hardware or computer* or device* or system* or technolog* or machine* or platform*)) 24,330
34 #33 AND #4 451
33 TS=(MRGB or MR-GB or MRIGB or MRI-GB or MRIFB or MRI-FB or MRFTB or MRF-TB or MRFTB or MRF-TB or MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx or FBx or FUSTB or FUS-TB or TB-FUS or "Fusion TB" or MRI-TRUS or MRI-TRUSB or MRI-TPB or COG-TB or TB-COG or CBx or TRUS-TB or "MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB") 1,351
32 #31 OR #30 OR #25 OR #18 3,620
31 TS=(("MRI stratified" or "magnetic resonance imaging stratified") NEAR/3 pathway*) 3
30 #29 AND #17 AND #4 2,016
29 #26 OR #27 OR #28 22,800
28 TS=((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR/3 (guid* or influenc* or direct* or align*)) 17,122
27 TS=(focal NEAR/2 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI)) 1,243
26 TS=(target* NEAR/4 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI)) 5,682
25 #24 AND #4 2,567
24 #23 OR #22 OR #21 OR #20 OR #19 6,484
23 TS=(focal NEAR/2 (biopsy or biopsie* or rebiopsy or rebiopsie*)) 666
22 TS=(target* NEAR/4 (biopsy or biopsie* or rebiopsy or rebiopsie*)) 4,437
21 TS=((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR/6 (prior or previous* or preced* or before* or earlier or first or initial*) NEAR/6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*)) 858
20 TS=((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR/6 prebiops*) 179
19 TS=((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) NEAR/6 (prior or previous* or preced* or before* or earlier or first or initial*) NEAR/6 (biopsy or biopsie*)) 963
18 #17 AND #14 AND #7 AND #4 1,832
17 #15 AND #16 28,427
16 TS=(biopsy or biopsie* or rebiopsy or rebiopsie*) 379,853
15 TS=(prostate* or prostatic) 336,855
14 #13 OR #12 OR #11 OR #10 OR #9 OR #8 2,737,360
13 TS=(software or hardware) 906,626
12 TS=(elastic or rigid or nonrigid) 624,588
11 TS=(registration*) 134,030
10 TS=(visual* NEAR/3 (estimat* or direct* or align* or guid* or influenc*)) 43,631
9 TS=cognitive* 514,118
8 TS=(fusion* or fuse* or fusing*) 589,649
7 #5 OR #6 757,071

6 TS=(mpMRI or mp-MRI or "mpMR imag*" or "mpMR scan*" or "mp-MR imag*" or "mp-MR scan*" or bpMRI or bp-MRI or "bpMR imag*" or "bpMR scan*" or "bp-MR imag*" or "bp-MR scan*") 2,175
5 TS=("magnetic resonance" or MRI or "MR imag*" or "MR scan*") 756,868
4 #1 OR #2 OR #3 332,891
3 TS=(((atypical NEAR/3 proliferation) or ASAP) and prostat*) 317
2 TS=(PCa or sPCa or csPCa or PrCa) 101,467
1 TS=((prostate* or prostatic or intraprostatic) NEAR/4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*)) 246,739

Key:

TS = topic tag; searches in title, abstract, author keywords and keywords plus fields

* = truncation

NEAR/3 = terms within three words of each other (any order)

On-going, unpublished or grey literature search strategies

ClinicalTrials.gov

https://clinicaltrials.gov/ct2/

Date searched: 23rd May 2022

Records retrieved: 572

Targeted search screen
1. 87 Studies found for: (prostate OR prostatic OR intraprostatic) AND (neoplasm OR cancer OR tumour OR tumor OR lesion OR nodule OR adenocarcinoma) [condition] | (biopsy OR biopsied OR rebiopsy OR rebiopsied) AND (MRI OR MR OR "magnetic resonance" OR biparametric OR multiparametric OR bpMRI OR mpMRI OR bp-MRI OR mp-MRI)[title]

2. 238 Studies found for: (prostate OR prostatic OR intraprostatic) AND (neoplasm OR cancer OR tumour OR tumor OR lesion OR nodule OR adenocarcinoma) [condition] | (biopsy OR biopsied OR rebiopsy OR rebiopsied) AND (MRI OR MR OR "magnetic resonance" OR biparametric OR multiparametric OR bpMRI OR mpMRI OR bp-MRI OR mp-MRI) [intervention]

3. 53 Studies found for: (prostate OR prostatic OR intraprostatic) AND (neoplasm OR cancer OR tumour OR tumor OR lesion OR nodule OR adenocarcinoma) [condition] | (biopsy OR biopsied OR rebiopsy OR rebiopsied) AND (targeted) [title]

4. 129 Studies found for: (prostate OR prostatic OR intraprostatic) AND (neoplasm OR cancer OR tumour OR tumor OR lesion OR nodule OR adenocarcinoma) [condition] | (biopsy OR biopsied OR rebiopsy OR rebiopsied) AND (targeted) [intervention]

Main search screen
5. 21 Studies found for: KOELIS OR "Fusion Bx" OR Biojet OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu OR ExactVu [other terms]

6. 44 Studies found for: Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR

ARTEMIS OR ProFuse OR "Mona Lisa" [other terms] | (prostate OR prostatic OR intraprostatic) [condition]

Conference Proceedings Citation Index – Science (CPCI-Science)

via Web of Science, Clarivate Analytics https://clarivate.com/

Date range: 1990 - present (CPCI-Science)

Date searched: 16th May 2022

See above under Science Citation Index for search strategy used. The number of records retrieved from CPCI-Science is not available as both Science Citation Index and CPCI-Science were searched together retrieving 3616 records in total from both databases.

EU Clinical Trials Register

via https://www.clinicaltrialsregister.eu/ctr-search/search

Search date: 15th June 2022

Records retrieved: 86

1. 68 result(s) found for: (prostate OR prostatic OR intraprostatic) AND (biopsy OR rebiopsy OR re-biopsy) AND (neoplasm OR cancer OR tumour OR tumor OR lesion OR nodule OR adenocarcinoma) AND (MRI OR MR OR "magnetic resonance" OR biparametric OR multiparametric OR bpMRI OR mpMRI OR bp-MRI OR mp-MRI) date range: 2015-01-01 to 2022-06-15

2. 18 result(s) found for: (prostate OR prostatic OR intraprostatic) AND (biopsy OR rebiopsy OR re-biopsy) AND (neoplasm OR cancer OR tumour OR tumor OR lesion OR nodule OR adenocarcinoma) AND targeted date range: 2015-01-01 to 2022-06-15

3. 0 result(s) found for: (KOELIS OR "Fusion Bx" OR Biojet OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu OR ExactVu) date range: 2015-01-01 to 2022-06-15

4. 0 result(s) found for: (Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa") AND (prostate OR prostatic OR intraprostatic) date range: 2015-01-01 to 2022-06-15

Open Access Theses and Dissertations (OATD)

via https://oatd.org/

Date searched: 16th May 2022

Records retrieved: 74

3 search queries used:

Query 1

(Prostat* AND biops*) AND (fusion* OR cognitive* OR software) AND (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*) AND ("magnetic resonance" OR MRI OR biparametric OR multiparametric)
50 hits

Query 2
(cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*) AND (prostat*) AND (Biojet OR Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa")
23 hits

Query 3
KOELIS OR "Fusion Bx" OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu* OR ExactVu*
1 hit

Key:

* = truncation

ProQuest Dissertations and Theses A&I

via https://www.proquest.com

Date searched: 16th May 2022

Records retrieved: 207

1. ((TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) OR TI,AB,SU,IF(((atypical NEAR/3 proliferation) OR ASAP) AND prostat*)) AND (TI,AB,SU,IF("magnetic resonance" OR MRI OR MR imag* OR MR scan*) OR TI,AB,SU,IF(mpMRI OR mp-MRI OR mpMR imag* OR mpMR scan* OR mp-MR imag* OR mp-MR scan* OR bpMRI OR bp-MRI OR bpMR imag* OR bpMR scan* OR bp-MR imag* OR bp-MR scan*)) AND (TI,AB,SU,IF(prostate* OR prostatic) AND TI,AB,SU,IF(biopsy OR biopsie* OR rebiopsy OR rebiopsie*)) AND (TI,AB,SU,IF(fusion* OR fuse* OR fusing* OR cognitive* OR registration* OR elastic OR rigid OR nonrigid OR software OR hardware) OR TI,AB,SU,IF(visual* NEAR/3 (estimat* OR direct* OR align* OR guid* OR influenc*))) limit: 2008-01-01 to 2022-05-16
33 Hits

2. (TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) AND TI,AB,SU,IF((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) NEAR/6 (biopsy OR biopsie* OR prebiops* OR ultrasound* OR ultrasonic* OR ultrasonograph* OR TRUS OR transperineal* OR transrectal*)) limit: 2008-01-01 to 2022-05-16 67 hits

3. (TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion*

OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) AND (TI,AB,SU,IF((target* OR focal) NEAR/4 (biopsy OR biopsie* OR rebiopsy OR rebiopsie*)) OR TI,AB,SU,IF((target* OR focal) NEAR/4 (MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI))) limit: 2008-01-01 to 2022-05-16 53 hits

4. ((TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) OR TI,AB,SU,IF(((atypical NEAR/3 proliferation) OR ASAP) AND prostat*)) AND (TI,AB,SU,IF(prostate* OR prostatic) AND TI,AB,SU,IF(biopsy OR biopsie* OR rebiopsy OR rebiopsie*)) AND TI,AB,SU,IF((MRI OR MR OR "magnetic resonance" OR mpMRI OR mp-MRI OR bpMRI OR bp-MRI) NEAR/3 (guid* OR influenc* OR direct* OR align*)) limit: 2008-01-01 to 2022-05-16 20 hits

5. ((TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) OR TI,AB,SU,IF(((atypical NEAR/3 proliferation) OR ASAP) AND prostat*)) AND TI,AB,SU,IF(MRGB OR MR-GB OR MRIGB OR MRI-GB OR MRIFB OR MRI-FB OR MRFTB OR MRF-TB OR MRFTB OR MRF-TB OR MRTB OR MR-TB OR MRITB OR MRI-TB OR MRTBx OR MR-TBx OR MRITBx OR MRI-TBx OR FBx OR FUSTB OR FUS-TB OR TB-FUS OR "Fusion TB" OR MRI-TRUS OR MRI-TRUSB OR MRI-TPB OR COG-TB OR TB-COG OR CBx OR TRUS-TB OR "MRI/TRUS" OR "mpMRI/TRUS" OR "MR/US" OR "MRI/TRUS-TB") limit: 2008-01-01 to 2022-05-16 6 hits

6. ((TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) OR TI,AB,SU,IF(((atypical NEAR/3 proliferation) OR ASAP) AND prostat*)) AND TI,AB,SU,IF(fusion* NEAR/3 (software OR hardware OR computer* OR device* OR system* OR technolog* OR machine* OR platform*)) limit: 2008-01-01 to 2022-05-16 26 hits

7. TI,AB,SU,IF(KOELIS OR "Fusion Bx" OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu* OR ExactVu*) 0 hits

8. ((TI,AB,SU,IF((prostate* OR prostatic OR intraprostatic) NEAR/4 (cancer* OR neoplas* OR tumour* OR tumor* OR malignan* OR metasta* OR carcinoma* OR adenocarcinoma* OR lesion* OR nodul* OR sarcoma* OR lymphoma*)) OR TI,AB,SU,IF(PCa OR sPCa OR csPCa OR PrCa)) OR TI,AB,SU,IF(((atypical NEAR/3 proliferation) OR ASAP) AND prostat*)) AND TI,AB,SU,IF(Biojet OR Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion") limit: 2008-01-01 to 2022-05-16 2 hits

Key:

TI,AB,SU,IF = search of title, abstract, subject heading and keyword fields

* = truncation

NEAR/3 = terms within three words of each other (any order)

PROSPERO

via https://www.crd.york.ac.uk/prospero/

Date searched: 23rd May 2022

Records retrieved: 78

#1 MeSH DESCRIPTOR Prostatic Intraepithelial Neoplasia 0
#2 MeSH DESCRIPTOR Prostatic Neoplasms EXPLODE ALL TREES 406
#3 (prostate* or prostatic or intraprostatic) adj4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*) 1351
#4 ((prostate* or prostatic or intraprostatic) adj4 (cancer* or neoplas* or tumour* or tumor* or malignan* or metasta* or carcinoma* or adenocarcinoma* or lesion* or nodul* or sarcoma* or lymphoma*)):TI 740
#5 ((prostate* or prostatic or intraprostatic)):TI 1080
#6 PCa or sPCa or csPCa or PrCa 335
#7 #1 OR #2 OR #4 OR #6 951
#8 #1 OR #2 OR #3 OR #6 1509
#9 #5 OR #2 OR #1 1092
#10 MeSH DESCRIPTOR Magnetic Resonance Imaging 458
#11 MeSH DESCRIPTOR Multiparametric Magnetic Resonance Imaging 6
#12 "magnetic resonance" or MRI or MR imag* or MR scan* 5234
#13 ("magnetic resonance" or MRI or MR imag* or MR scan*):TI 773
#14 ((mpMRI or mp-MRI or mpMR imag* or mpMR scan* or mp-MR imag* or mp-MR scan* or bpMRI or bp-MRI or bpMR imag* or bpMR scan* or bp-MR imag* or bp-MR scan*)):TI 8
#15 (mpMRI or mp-MRI or mpMR imag* or mpMR scan* or mp-MR imag* or mp-MR scan* or bpMRI or bp-MRI or bpMR imag* or bpMR scan* or bp-MR imag* or bp-MR scan*) 59
#16 #10 OR #11 OR #12 OR #15 5259
#17 #10 OR #11 OR #13 OR #14 887
#18 MeSH DESCRIPTOR Image Interpretation, Computer-Assisted 4
#19 MeSH DESCRIPTOR Software 31
#20 fusion* or fuse* or fusing* or cognitive or registration* or elastic or rigid or nonrigid 17958
#21 visual* adj3 (estimat* or direct* or align* or guid* or influenc*) 274
#22 software or hardware 48745
#23 #18 OR #19 OR #20 OR #21 OR #22 60890
#24 MeSH DESCRIPTOR Prostate 102
#25 prostate* or prostatic 1862
#26 (prostate* or prostatic):TI 1080
#27 #24 OR #25 1881
#28 #24 OR #26 1102
#29 (MeSH DESCRIPTOR Biopsy):TI 0
#30 (MeSH DESCRIPTOR Image-Guided Biopsy EXPLODE ALL TREES):TI 0
#31 (MeSH DESCRIPTOR Biopsy, Needle EXPLODE ALL TREES):TI 0
#32 MeSH DESCRIPTOR Biopsy, Needle EXPLODE ALL TREES 50
#33 MeSH DESCRIPTOR Biopsy, Needle EXPLODE ALL TREES 50
#34 MeSH DESCRIPTOR Biopsy 103
#35 MeSH DESCRIPTOR Image-Guided Biopsy EXPLODE ALL TREES 27
#36 biopsy or biopsie* or rebiopsy or rebiopsie* 2655
#37 (biopsy or biopsie* or rebiopsy or rebiopsie*):TI 251
#38 #32 OR #34 OR #35 OR #36 2678
#39 #32 OR #34 OR #35 OR #37 295
#40 #8 AND #16 AND #23 AND #27 AND #38 54
#41 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous* or preced* or before* or earlier or first or initial*) adj6 (biopsy or biopsie*)):TI 1

#42 ((MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 prebiops*):TI 0
#43 ((MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous* or preced* or before* or earlier or first or initial*) adj6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*)):TI 0
#44 (target* adj4 (biopsy or biopsie* or rebiopsy or rebiopsie*)):TI 15
#45 (focal* adj2 (biopsy or biopsie* or rebiopsy or rebiopsie*)):TI 1
#46 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous* or preced* or before* or earlier or first or initial*) adj6 (biopsy or biopsie*) 9
#47 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 prebiops* 0
#48 (MRI or MR or magnetic resonance or mpMRI or mp-MRI or bpMRI or bp-MRI) adj6 (prior or previous* or preced* or before* or earlier or first or initial*) adj6 (ultrasound* or ultrasonic* or ultrasonograph* or TRUS or transperineal* or transrectal*) 0
#49 target* adj4 (biopsy or biopsie* or rebiopsy or rebiopsie*) 48
#50 focal* adj2 (biopsy or biopsie* or rebiopsy or rebiopsie*) 1
#51 #46 OR #47 OR #48 OR #49 OR #50 55
#52 #8 AND #51 44
#53 #52 OR #40 66
#54 target* adj4 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) 22
#55 focal* adj2 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) 2
#56 (MRI or MR or "magnetic resonance" or mpMRI or mp-MRI or bpMRI or bp-MRI) adj3 (guid* or influenc* or direct* or align*) 139
#57 #54 OR #55 OR #56 154
#58 #8 AND #27 AND #38 AND #57 38
#59 #53 OR #58 76
#60 MRGB or MR-GB or MRIGB or MRI-GB or MRIFB or MRI-FB or MRFTB or MRF-TB or MRFTB or MRF-TB or MRTB or MR-TB or MRITB or MRI-TB or MRTBx or MR-TBx or MRITBx or MRI-TBx or FBx or FUSTB or FUS-TB or TB-FUS or "Fusion TB" or MRI-TRUS or MRI-TRUSB or MRI-TPB or COG-TB or TB-COG or CBx or TRUS-TB or "MRI/TRUS" or "mpMRI/TRUS" or "MR/US" or "MRI/TRUS-TB" 17
#61 fusion* adj3 (software or hardware or computer* or device* or system* or technolog* or machine* or platform*) 75
#62 #60 OR #61 91
#63 #8 AND #62 16
#64 #63 OR #59 76
#65 KOELIS OR "Fusion Bx" OR Biojet OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu* OR ExactVu* 7
#66 Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa" 489
#67 #66 AND #8 4
#68 #64 OR #65 OR #67 78

Key:

MeSH DESCRIPTOR = subject heading (MeSH heading)

* = truncation

adj3 = terms within 3 words of each other (order specified)

WHO International Clinical Trials Registry Platform (ICTRP)

https://trialsearch.who.int/AdvSearch.aspx

Date searched: 23rd May 2022

Records retrieved: 378

Advanced search screen. Recruitment status set to ALL
1. Title field: (biops* OR rebiops* OR re-biops*) AND (MRI OR MR OR "magnetic resonance" OR biparametric OR multiparametric OR bpMRI OR mpMRI OR bp-MRI OR mp-MRI)
Condition field: (prostate* OR prostatic OR intraprostatic) AND (neoplas* OR cancer* OR tumour* OR tumor* OR lesion* OR nodul* OR adenocarcinoma*)
117 hits

2. Intervention field: (biops* OR rebiops* OR re-biops*) AND (MRI OR MR OR "magnetic resonance" OR biparametric OR multiparametric OR bpMRI OR mpMRI OR bp-MRI OR mp-MRI)
Condition field: (prostate* OR prostatic OR intraprostatic) AND (neoplas* OR cancer* OR tumour* OR tumor* OR lesion* OR nodul* OR adenocarcinoma*)
106 hits

3. Title field: (biops* OR rebiops* OR re-biops*) AND target*
Condition field: (prostate* OR prostatic OR intraprostatic) AND (neoplas* OR cancer* OR tumour* OR tumor* OR lesion* OR nodul* OR adenocarcinoma*)
68 hits

4. Intervention field: (biops* OR rebiops* OR re-biops*) AND target*
Condition field: (prostate* OR prostatic OR intraprostatic) AND (neoplas* OR cancer* OR tumour* OR tumor* OR lesion* OR nodul* OR adenocarcinoma*)
64 hits

5. Title field: (Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa")
Condition field: (prostate* OR prostatic OR intraprostatic) AND (neoplas* OR cancer* OR tumour* OR tumor* OR lesion* OR nodul* OR adenocarcinoma*)
4 hits

6. Intervention field: (Trinity OR PROMAP OR "Fusion MR" OR bkFusion OR "bk Fusion" OR BK3000 OR "BK 3000" OR BK5000 OR "BK 5000" OR "Predictive Fusion" OR DynaCAD OR ARTEMIS OR ProFuse OR "Mona Lisa")
Condition field: (prostate* OR prostatic OR intraprostatic) AND (neoplas* OR cancer* OR tumour* OR tumor* OR lesion* OR nodul* OR adenocarcinoma*)
3 hits

Basic search screen

7. KOELIS OR "Fusion Bx" OR Biojet OR Biopsee OR UroNav OR "iSR'obot" OR iSRobot OR "iSR obot" OR UroFusion OR UroBiopsy OR FusionVu* OR ExactVu*

16 hits

Key:

* = truncation

Guideline website searches

Simple searches were carried out on the guideline websites listed below and any results were browsed for relevance. Relevant guidelines identified were checked against the endnote library of results and added to the library if they had not already been found through previous searches.

ECRI Guidelines Trust

https://guidelines.ecri.org/

Date searched: 23rd May 2022

1. prostate or prostatic - 39 results browsed – 9 relevant

GIN international guideline library

https://guidelines.ebmportal.com/

Date searched: 23rd May 2022

1. prostate cancer - 36 results browsed – 8 relevant

National Institute of health and Care Excellence (NICE)

https://www.nice.org.uk/

Date searched: 23rd May 2022

1. Browsed 43 items on the prostate cancer guidance page
https://www.nice.org.uk/guidance/conditions-and-diseases/cancer/prostate-cancer

- 4 relevant

Trip database

https://www.tripdatabase.com/

Date searched: 23rd May 2022

2 further guidelines found through searching the Trip database.

1. Prostate cancer AND MRI OR "magnetic resonance" OR biparametric OR multiparametric - 5 guideline results – browsed for relevance – 4 relevant – all in endnote library already.

2. Prostate cancer AND fusion OR cognitive OR software – 0 guideline results

3. Prostate cancer AND imag* - 6 guideline results - browsed for relevance – 3 relevant – all in endnote library already.

4. Prostate cancer AND diagnos* - 10 guideline results – browsed for relevance - 8 relevant – 6 in endnote library already.

APPENDIX 2. MULTINOMIAL NMA MODEL DESCRIPTION

A multinomial logistic regression model was used where the odds of being categorised in each of the different categories in Table 3 compared to the reference category (no PCa) are allowed to vary by biopsy type.[74, 75, 168, 169] This model is conceptually equivalent to R-1 binomial logistic regressions comparing category $r>1$ with category 1 (no PCa), for each different biopsy type compared to the reference, cognitive biopsy.

Define

i – study index

k – study arm index

r – category index

R – number of categories

Data from the N studies are modelled with a multinomial likelihood with probability vector q_{ikr}

$$y_{ik,1:R} \sim \text{Multinomial}\left(q_{ik,1:R}, M_{ik}\right)$$

$y_{ik,1:R}$ - vector of observed events in arm k of study i

M_{ik} - number of patients in arm k of study i

Category probabilities for arm k of study i are defined as

$$q_{ik,r} = \frac{\phi_{ikr}}{\sum_{s=1}^{R} \phi_{iks}}$$

Log-odds ratio for category r relative to category 1, for arm k in study i:

$$\eta_{ikr} = \log\left(\frac{q_{ikr}}{q_{ik1}}\right) = a_{ir} + \delta_{ikr} \quad (1)$$

with a_{ir} representing the baseline log-odds for being classified in category r, instead of category 1, in study i and $\delta_{ikr} = d_{t_{i1}t_{ik},r} = d_{1t_{ik},r} - d_{1t_{i1},r}$ representing the additional effect for being classified in

category r, instead of category 1, using the intervention in arm k, compared to the intervention in arm 1.

We set

$$d_{1r} = 0, \text{ for all } r$$
$$d_{k1} = 0, \text{ for all } k$$
$$a_{i1} = 0, \text{ for all } i$$

Note that

$$q_{ik1} = \frac{\phi_{ik1}}{\sum_{s=1}^{R} \phi_{iks}} = \frac{1}{\sum_{s=1}^{R} \phi_{iks}}$$

Hence

$$\begin{aligned} \phi_{ikr} &= q_{ikr} \times \sum_{s=1}^{R} \phi_{iks} \\ &= q_{ikr} \times \frac{1}{q_{ik1}} = \frac{q_{ikr}}{q_{ik1}} \\ &= \exp\left(\eta_{ikr}\right) = \exp\left(a_{ir} + \delta_{ikr}\right) \end{aligned}$$

We model ϕ_{ikr}, the odds ratio for category r relative to category 1, for arm k in study i as

$$\log\left(\phi_{ikr}\right) = a_{ir} + \delta_{ikr} \tag{2}$$

Calculating absolute probabilities

To calculate the absolute probabilities of being classified in category r using **intervention** k, T_{kr} we note:

$$\begin{aligned} T_{kr} &= \frac{\phi_{kr}}{\sum_{s=1}^{R} \phi_{ks}} \\ T_{k1} &= \frac{1}{\sum_{s=1}^{R} \phi_{ks}} \end{aligned} \tag{3}$$

Using equation (2), and defining A_r as the log-odds of being classified in category r using the reference intervention, we have

$$\begin{aligned} \log\left(\phi_{k1}\right) &= A_1 + d_{k1} = 0 \\ \phi_{k1} &= 1 \end{aligned} \tag{4}$$

and using equation (1) we have

$$\begin{aligned} \log\left(\frac{T_{kr}}{T_{k1}}\right) &= A_r + d_{kr} \\ \log\left(T_{kr}\right) &= \log\left(T_{k1}\right) + A_r + d_{kr} \\ T_{kr} &= \exp\left(\log\left(T_{k1}\right) + A_r + d_{kr}\right) \end{aligned}$$

External data inform T_{k1} which are used to calculate A_r and calculate all the other probabilities

Using equations (3) and (4) we have

$$\begin{aligned} \log\left(\phi_{kr}\right) &= A_r + d_{kr} \\ T_{k1} &= \frac{1}{1 + \phi_{k2} + \ldots + \phi_{kR}} \end{aligned}$$

WinBUGS code for multinomial model

Code

```
model{
for (i in 1:ns){ # studies reporting all categories 1,2,3,4,5
  for (k in 1:na[i]) {
    y[i,k,1:nc] ~ dmulti(q[i,k,1:nc], M[i,k])
    for (r in 1:nc) {
      q[i,k,r] <- phi[i,k,r]/sum(phi[i,k,])
      log(phi[i,k,r]) <- a[i,r] + d[t[i,k],r] - d[t[i,1],r]
     # predicted number events
     yhat[i,k,r] <-  q[i,k,r] * M[i,k]
     # Deviance contribution
     dv[i,k,r] <- 2*y[i,k,r]*(log(y[i,k,r]/yhat[i,k,r]))
     }
    dev[i,k] <- sum(dv[i,k,]) # deviance contribution of each arm
   }
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  for (r in 2:nc) {a[i,r] ~ dnorm(0, 0.0001)}
  a[i,1] <- 0
  # summed residual deviance contribution for this trial
  resdev[i] <- sum(dev[i,1:na[i]])
 }
#relative effects of treatment 1 compared to itself are zero, for all categories
for (r in 2:nc) {d[1,r] <- 0}
for (k in 1:nt){
  # giving phi[i,k,1] = 1, logOR of  going from cat 1 to cat 1 for all treats
  d[k,1] <- 0
  for (r in 2:nc) {
```

```
    # vague priors for relative treatment effects: log-odds ratios
    d[k,r] ~ dnorm(0, 0.0001)
   }
 }
# STUDIES WITH COLLAPSED CATEGORIES: TYPE A
for (i in (ns+1):(ns+nsA)){ # studies reporting categories 1,2-5
  for (k in 1:na[i]) {
    y[i,k,1] ~ dbin(q[i,k,1], M[i,k])
    # first category the same
    q[i,k,1] <- phi[i,k,1]/sum(phi[i,k,1:ncA])
    log(phi[i,k,1]) <- a[i,1] + d[t[i,k],1] - d[t[i,1],1]
     yhat[i,k,1] <-  q[i,k,1] * M[i,k]
     # Deviance contribution
     dev[i,k] <- 2 * (y[i,k,1] * (log(y[i,k,1])-log(yhat[i,k,1])) +  (M[i,k]-
y[i,k,1]) * (log(M[i,k]-y[i,k,1]) - log(M[i,k]-yhat[i,k,1])))
# last category is collapsed, type A
    q[i,k,2] <- 1- q[i,k,1]
    log(phi[i,k,2]) <- a[i,2] + dA[t[i,k],2] - dA[t[i,1],2]
   }
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  a[i,2] ~ dnorm(0, 0.0001)
  a[i,1] <- 0
 # summed residual deviance contribution for this trial
 resdev[i] <- sum(dev[i,1:na[i]])
 }
dA[1,2] <- 0
for (k in 1:nt){
  # vague priors for relative treatment effects: log-odds ratios
  dA[k,2] ~ dnorm(0, 0.0001)
   }
# STUDIES WITH COLLAPSED CATEGORIES: TYPE B
for (i in (ns+nsA+1):(ns+nsA+nsB)){ # studies reporting categories 1,2,3-5
  for (k in 1:na[i]) {
    y[i,k,1:ncB] ~ dmulti(q[i,k,1:ncB], M[i,k])
    for (r in 1:(ncB-1)) {  # first 2 categories the same
      q[i,k,r] <- phi[i,k,r]/sum(phi[i,k,1:ncB])
      log(phi[i,k,r]) <- a[i,r] + d[t[i,k],r] - d[t[i,1],r]
     # predicted number events
     yhat[i,k,r] <-  q[i,k,r] * M[i,k]
     # Deviance contribution
     dv[i,k,r] <- 2*y[i,k,r]*(log(y[i,k,r]/yhat[i,k,r]))
     }
# last category is collapsed, type B
    q[i,k,3] <- 1- sum(q[i,k,1:(ncB-1)])
    log(phi[i,k,3]) <- a[i,3] + dB[t[i,k],3] - dB[t[i,1],3]
   # predicted number events
   yhat[i,k,3] <-  q[i,k,3] * M[i,k]
   # Deviance contribution
    dv[i,k,3] <- 2*y[i,k,3]*(log(y[i,k,3]/yhat[i,k,3]))
    dev[i,k] <- sum(dv[i,k,1:ncB]) # deviance contribution of each arm
   }
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  for (r in 2:ncB) {a[i,r] ~ dnorm(0, 0.0001)}
  a[i,1] <- 0
  # summed residual deviance contribution for this trial
  resdev[i] <- sum(dev[i,1:na[i]])
 }
dB[1,ncB] <- 0
for (k in 1:nt){
  # vague priors for relative treatment effects: log-odds ratios
  dB[k,ncB] ~ dnorm(0, 0.0001)
   }
# STUDIES WITH COLLAPSED CATEGORIES: TYPE C
```

```
for (i in (ns+nsA+nsB+1):(ns+nsA+nsB+nsC)){ # studies reporting categories 1,2,3,4-
5
  for (k in 1:na[i]) {
    y[i,k,1:ncC] ~ dmulti(q[i,k,1:ncC], M[i,k])
    for (r in 1:(ncC-1)) {  # first 3 categories the same
      q[i,k,r] <- phi[i,k,r]/sum(phi[i,k,1:ncC])
      log(phi[i,k,r]) <- a[i,r] + d[t[i,k],r] - d[t[i,1],r]
     # predicted number events
     yhat[i,k,r] <-  q[i,k,r] * M[i,k]
     # Deviance contribution
     dv[i,k,r] <- 2*y[i,k,r]*(log(y[i,k,r]/yhat[i,k,r]))
     }
# last category is collapsed, type C
    q[i,k,4] <- 1- sum(q[i,k,1:(ncC-1)])
    log(phi[i,k,4]) <- a[i,4] + dC[t[i,k],4] - dC[t[i,1],4]
   # predicted number events
   yhat[i,k,4] <-  q[i,k,4] * M[i,k]
   # Deviance contribution
   dv[i,k,4] <- 2*y[i,k,4]*(log(y[i,k,4]/yhat[i,k,4]))
   dev[i,k] <- sum(dv[i,k,1:ncC]) # deviance contribution of each arm
   }
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  for (r in 2:ncC) {a[i,r] ~ dnorm(0, 0.0001)}
  a[i,1] <- 0
  # summed residual deviance contribution for this trial
  resdev[i] <- sum(dev[i,1:na[i]])
 }
dC[1,ncC] <- 0
for (k in 1:nt){
  # vague priors for relative treatment effects: log-odds ratios
  dC[k,ncC] ~ dnorm(0, 0.0001)
   }
totresdev <- sum(resdev[])        # Total Residual Deviance
# pairwise ORs and LORs for all possible pair-wise comparisons
for (c in 1:(nt-1)) {
   for (k in (c+1):nt)  {
      for (r in 1:nc) {
        or[c,k,r] <- exp(d[k,r] - d[c,r])
        lor[c,k,r] <- (d[k,r]-d[c,r])
         }
     }
 }
# calculate absolute probabilities from relative effects (no uncertainty in
baseline)
# baseline intervention = 3 (software) or 5 (software + SB)
for (r in 1:nc){
  T[3,r] <- b1[r]  # baseline probabilities for software biopsy (from data)
  T[5,r] <- b2[r]  # baseline probabilities for software + SB (from data)
  # log-odds of being classified in category r using intervention 3=software
  A.T[r] <- log(T[3,r]/T[3,1])
  # log-odds of being classified in category r using intervention 5=software +SB
  B.T[r] <- log(T[5,r]/T[5,1])
 }
for (k in 1:2){                  # fully connected only T[2,] to T[3,]
  for (r in 1:nc){
    phi.T[k,r]  <- exp(A.T[r] - lor[k,3,r])
    T[k,r] <- phi.T[k,r]/(sum(phi.T[k,]))
   }
 }
for (r in 1:nc){
  phi.T[4,r]  <- exp(B.T[r] - lor[4,5,r])
  T[4,r] <- phi.T[4,r]/(sum(phi.T[4,]))
 }
}
```

Data

```
# ns = number of studies
# nt = number of treatments
# nc = number of categories
# nsX = number of studies of type X
# ncX = number of categories in studies type X

# T1=cog
# T2=SB
# T3=fus
# T4=cog+SB
# T5=fus+SB

list(ns=2, nt=5, nc=5, nsA=4, ncA=2, nsB=5, ncB=3, nsC=2, ncC=4,
#b1=c(0.379032,0.153226,0.209677,0.157258,0.100806),  # PAIREDCAP baseline probs - cognitive
#b1=c(0.286290,0.173387,0.282258,0.161290,0.096774),  # PAIREDCAP baseline probs - software (Artemis)
#b1=c(0.468864,0.164835,0.197802,0.105311,0.063187),  # Filson (naive only) baseline probs - software (Artemis)
b1=c(0.686792,0.086792,0.101887,0.077830,0.046698),  # Filson (prior neg) baseline probs - software (Artemis)
#b2=c(0.355311,0.219780,0.223443,0.118321,0.083144)  # Filson (naive only) baseline probs - software + SB (Artemis) [split by SB proportion in PAIREDCAP]
b2=c(0.584906,0.150943,0.116981,0.086433,0.060737)  # Filson (prior neg) baseline probs - software + SB (Artemis) [split by SB proportion in PAIREDCAP]
#b2=c(0.355311,0.219780,0.223443,0.125916,0.075549)  # Filson (naive only) baseline probs - software + SB (Artemis) [split by Artemis proportion in PAIREDCAP]
)
```

na[]	t[,1]	t[,2]	t[,3]	M[,1]	M[,2]	M[,3]	y[,1,1]	y[,2,1]	y[,3,1]	y[,1,2]	y[,2,2]	
	y[,3,2]	y[,1,3]	y[,2,3]	y[,3,3]	y[,1,4]	y[,2,4]	y[,3,4]	y[,1,5]	y[,2,5]	y[,3,5]	#	study
ID	Study type											
3	1	2	3	248	248	248	94	52	71	38	46	43
	52	87	70	39	37	40	25	26	24	#	PAIREDCAP	
	all											
2	4	5	NA	100	99	NA	69	55	NA	19	25	NA
	6	13	NA	5	3	NA	1	3	NA	#	Izadpanahi	
	all											
3	2	3	5	169	169	169	53	49	36	116	120	133
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Wajswol 2020	
	A											
3	1	2	4	75	75	75	41	35	32	34	40	43
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Thangarasu	
2021	A											
3	1	2	4	63	63	63	30	33	25	33	30	38
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Kulis 2020	
	A											
2	1	3	NA	88	88	NA	57	48	NA	31	40	NA
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Cornud	A
2	1	3	NA	78	79	NA	44	40	NA	8	12	NA
	26	27	NA	NA	NA	NA	NA	NA	NA	#	FUTURE	B
2	1	3	NA	125	125	NA	85	80	NA	16	16	NA
	24	29	NA	NA	NA	NA	NA	NA	NA	#	PROFUS	B
3	2	3	5	74	74	74	41	39	32	12	10	13
	21	25	29	NA	NA	NA	NA	NA	NA	#	Albisinni 2018	
	B											
3	2	3	5	191	191	191	103	106	85	36	25	34
	52	60	72	NA	NA	NA	NA	NA	NA	#	Fourcade 2018	
	B											
3	1	2	4	111	111	111	69	81	65	19	9	20
	23	21	26	NA	NA	NA	NA	NA	NA	#	Gomez-Ortiz	
2022	B											
3	2	3	5	48	48	48	23	20	16	11	11	13
	10	13	13	4	4	6	NA	NA	NA	#	Alberts 2018	
(all men)	C											
3	2	3	5	538	538	538	294	310	252	114	68	100
	74	81	92	56	79	94	NA	NA	NA	#	Filson 2016	
	C											
END												

Initial values

```
#chain 1
list( a = structure(.Data = c(  NA,0,0,0,0,          NA,0,0,0,0,          NA,0,NA,NA,NA,          NA,0,NA,NA,NA,
NA,0,NA,NA,NA,          NA,0,NA,NA,NA,          NA,0,0,NA,NA,          NA,0,0,NA,NA,
NA,0,0,NA,NA,          NA,0,0,NA,NA,          NA,0,0,0,NA,          NA,0,0,0,NA),
.Dim = c(13,5)),
d = structure(.Data = c( NA,NA,NA,NA,NA,          NA,0,0,0,0,          NA,0,0,0,0,          NA,0,0,0,0,
NA,0,0,0,0),
.Dim = c(5,5)),
dA = structure(.Data = c(NA,NA,          NA,0,          NA,0,          NA,0,          NA,0),
.Dim = c(5,2)),
dB = structure(.Data = c(NA,NA,NA,          NA,NA,0,          NA,NA,0,          NA,NA,0,          NA,NA,0),
.Dim = c(5,3)),
dC = structure(.Data = c(NA,NA,NA,NA,          NA,NA,NA,0,          NA,NA,NA,0,          NA,NA,NA,0,
NA,NA,NA,0),
.Dim = c(5,4))
)

#chain 2
list( a = structure(.Data = c(  NA,2,-.5,1,-1,          NA,2,3,1,2,          NA,-2,NA,NA,NA,          NA,-2,NA,NA,NA,
NA,-2,NA,NA,NA,          NA,-2,NA,NA,NA,          NA,-2,1,NA,NA,          NA,1,-2,NA,NA,          NA,-2,1,NA,NA,
NA,-2,1,NA,NA,          NA,-2,1,NA,NA,          NA,.7,-2,-1,NA,          NA,1,-2,2,NA),
.Dim = c(13,5)),
d = structure(.Data = c( NA,NA,NA,NA,NA,          NA,-1,-2,1,2,          NA,.5,-2,-1,1,          NA,2,-2,.5,-2,
NA,1,2,1,-2),
.Dim = c(5,5)),
dA = structure(.Data = c(NA,NA,          NA,-2,          NA,2,          NA,1,          NA,2),
.Dim = c(5,2)),
dB = structure(.Data = c(NA,NA,NA,          NA,NA,1,          NA,NA,-1,          NA,NA,-2,          NA,NA,-1),
.Dim = c(5,3)),
dC = structure(.Data = c(NA,NA,NA,NA,          NA,NA,NA,2,          NA,NA,NA,.7,          NA,NA,NA,-.5,
NA,NA,NA,-2),
.Dim = c(5,4))
)
```

APPENDIX 3. STUDIES EXCLUDED FROM THE SYSTEMATIC REVIEW OF CLINICAL EVIDENCE

Table 59 Studies excluded from the systematic review with reasons at full text stage

Study	Reason for exclusion
Abouassaly R, Klein EA, El-Shefai A, Stephenson A. Impact of using 29 MHz high-resolution micro-ultrasound in real-time targeting of transrectal prostate biopsies: initial experience. World J Urol. 2020;38(5):1201-6.	Wrong population (MicroUS is not standard practice)
Ahdoot M, Lebastchi AH, Long L, Wilbur AR, Gomella PT, Mehralivand S, et al. Using Prostate Imaging-Reporting and Data System (PI-RADS) scores to select an optimal prostate biopsy method: a secondary analysis of the trio study. Eur Urol Oncol. 2022;5(2):176-86.	SF vs SB, insufficient data for inclusion in indirect comparison (lack of separable data excluding prior positives)
Ahdoot M, Wilbur AR, Reese SE, Lebastchi AH, Mehralivand S, Gomella PT, et al. MRI-targeted, systematic, and combined biopsy for prostate cancer diagnosis. N Engl J Med. 2020;382(10):917-28.	SF vs SB, insufficient data for inclusion in indirect comparison.
Al Hussein Al Awamlh B, Marks LS, Sonn GA, Natarajan S, Fan RE, Gross MD, et al. Multicenter analysis of clinical and MRI characteristics associated with detecting clinically significant prostate cancer in PI-RADS (v2.0) category 3 lesions. Urol Oncol. 2020;38(7):637.e9-.e15.	Wrong study design

Alqahtani S, Zhang X, Wei C, Zhang Y, Szewczyk-Bieda M, Wilson J, et al. Predicting the performance of concurrent systematic random biopsies during image fusion targeted sampling of multi-parametric MRI detected prostate cancer. a prospective study (PRESET study). [published online ahead of print, Dec 21 2021]. Cancers. 2021;11.	Wrong intervention: out of scope SF
Altok M, Demirel C, Kang HC, Choi H, John D, Inguillo IA, et al. Impact of MRI/US fusion-guided prostate biopsy on biopsy-naive patients: a single urologist's experience. BJUI Compass. 2022;3(1):19-25.	Wrong study design
Andras I, Crisan D, Cata E, Tamas-Szora A, Caraiani C, Coman RT, et al. MRI-TRUS fusion guided prostate biopsy - initial experience and assessment of the role of contralateral lobe SB. Med Ultrason. 2019;21(1):37-44.	Wrong intervention: out of scope SF
Arsov C, Quentin M, Rabenalt R, Antoch G, Albers P, Blondin D. Repeat transrectal ultrasound biopsies with additional targeted cores according to results of functional prostate MRI detects high-risk prostate cancer in patients with previous negative biopsy and increased PSA - a pilot study. Anticancer Res 2012;32:1087-92.	Wrong outcome
Arsov C, Rabenalt R, Blondin D, Quentin M, Hiester A, Godehardt E, et al. Prospective randomized trial comparing magnetic resonance imaging (MRI)-guided in-bore biopsy to MRI-ultrasound fusion and transrectal ultrasound-guided prostate biopsy in patients with prior negative biopsies. Eur Urol. 2015;68(4):713-20.	Wrong population
Arsov C, Rabenalt R, Quentin M, Hiester A, Blondin D, Albers P, et al. Comparison of patient comfort between MR-guided in-bore and MRI/ultrasound fusion-guided prostate biopsies within a prospective randomized trial. *World J Urol* 2016;**34**:215-20.	Wrong comparator
Avolio PP, Lughezzani G, Paciotti M, Maffei D, Uleri A, Frego N, et al. The use of 29 MHz transrectal micro-ultrasound to stratify the prostate cancer risk in patients with PI-RADS III lesions at multiparametric MRI: A single institutional analysis. Urol Oncol 2021;39:832.e1-.e7.	Wrong comparator: MicroUS is not standard practice
Baco E, Rud E, Eri LM, Moen G, Vlatkovic L, Svindland A, et al. A randomized controlled trial to assess and compare the outcomes of two-core prostate biopsy guided by fused magnetic resonance and transrectal ultrasound images and traditional 12-core systematic biopsy. Eur Urol 2016;69:149-56.	Wrong outcome
Bae JH, Kim SH. Transrectal ultrasound-guided prostate biopsy versus combined magnetic resonance imaging-ultrasound fusion and SB for prostate cancer detection in routine clinical practice. Ultrasonography. 2020;39(2):137-43.	Wrong study design
Ball MW, Ross AE, Ghabili K, Kim C, Jun C, Petrisor D, et al. Safety and feasibility of direct magnetic resonance imaging-guided transperineal prostate biopsy using a novel magnetic resonance imaging-safe robotic device. Urology. 2017;109:216-21.	Wrong intervention
Bansal S, Gupta NP, Yadav R, Khera R, Ahlawat K, Gautam D, et al. Multiparametric magnetic resonance imaging-transrectal ultrasound fusion prostate biopsy: a prospective, single centre study. Indian J Urol. 2017;33(2):134-9.	SF vs SB, insufficient data for inclusion in indirect comparison.
Barrett T, Patterson AJ, Koo BC, Wadhwa K, Warren AY, Doble A, et al. Targeted transperineal biopsy of the prostate has limited additional benefit over background cores for larger MRI-identified tumors. World J Urol. 2016;34(4):501-8.	Wrong study design
Bass EJ, Donaldson IA, Freeman A, Jameson C, Punwani S, Moore C, et al. Magnetic resonance imaging targeted transperineal prostate biopsy: a local anaesthetic approach. Prostate Cancer Prostatic Dis. 2017;20(3):311-7.	Wrong study design
Belas O, Klap J, Cornud F, Beuvon F, Peyromaure M, Zerbib M, et al. [Prebiopsy multiparametric MRI of the prostate: the end of randomized biopsies?]. Prog Urol. 2012;22(10):583-9.	Wrong study design
Benelli A, Vaccaro C, Guzzo S, Nedbal C, Varca V, Gregori A. The role of MRI/TRUS fusion biopsy in the diagnosis of clinically significant prostate cancer. [published online ahead of print May 18 2020]. Ther Adv Urol. 2020;12.	Wrong study design
Ber Y, Segal N, Tamir S, Benjaminov O, Yakimov M, Sela S, et al. A noninferiority within-person study comparing the accuracy of transperineal to transrectal MRI-US fusion biopsy for prostate-cancer detection. Prostate Cancer Prostatic Dis. 2020;23(3):449-56.	Wrong intervention: out of scope SF
Berkenwald A, Stensland KD, Sebel LE, Moinzadeh A, Faust W. Initial transperineal prostate biopsy experience at a high-volume center. Can J Urol. 2021;28(3):10692-8.	Wrong study design
Bhambri K, Pandey AK, Jhobta A, Bhambri A, Sharma S, Singh B, et al. Role of TRUS and MRI in the detection of prostate cancer-a prospective study. *J Clin Diagn Res* 2020;**14**:TC10-4.	Wrong comparator
Bladou F, Fogaing C, Levental M, Aronson S, Alameldin M, Anidjar M. Transrectal ultrasound-guided biopsy for prostate cancer detection: systematic and/or magnetic-resonance imaging-targeted. Can Urol Assoc J. 2017;11(9):E330-7.	SF and CF combined, no separate data per fusion method

Boesen L, Noergaard N, Chabanova E, Logager V, Balslev I, Mikines K, et al. Early experience with multiparametric magnetic resonance imaging-targeted biopsies under visual transrectal ultrasound guidance in patients suspicious for prostate cancer undergoing repeated biopsy. Scand J Urol 2015;49:25-34.	Wrong outcome
Boesen L, Norgaard N, Logager V, Balslev I, Bisbjerg R, Thestrup KC, et al. Assessment of the diagnostic accuracy of biparametric magnetic resonance imaging for prostate cancer in biopsy-naive men: the Biparametric MRI for Detection of Prostate Cancer (BIDOC) study. JAMA Netw Open 2018;1:e180219.	Wrong outcome
Boesen L, Norgaard N, Logager V, Balslev I, Thomsen HS. A prospective comparison of selective multiparametric magnetic resonance imaging fusion-targeted and systematic transrectal ultrasound-guided biopsies for detecting prostate cancer in men undergoing repeated biopsies. Urol Int. 2017;99(4):384-91.	Wrong intervention: out of scope SF
Boesen L, Norgaard N, Logager V, Balslev I, Thomsen HS. Multiparametric MRI in men with clinical suspicion of prostate cancer undergoing repeat biopsy: a prospective comparison with clinical findings and histopathology. Acta Radiol. 2018;59(3):371-80.	SF and CF combined, no separate data per fusion method
Boesen L, Norgaard N, Logager V, Balslev I, Thomsen HS. Where do transrectal ultrasound- and magnetic resonance imaging-guided biopsies miss significant prostate cancer? Urology. 2017;110:154-60.	Wrong study design
Boesen L, Norgaard N, Logager V, Thomsen HS. Clinical outcome following low suspicion multiparametric prostate magnetic resonance imaging or benign magnetic resonance imaging guided biopsy to detect prostate cancer. J Urol. 2017;198(2):310-5.	SF and CF combined, no separate data per fusion method
Bonet X, Suarez-Novo JF, Castells M, Serrallach M, Beato S, Picola N, et al. [Targeted biopsies using magnetic resonance imaging/ultrasonograpgy fusion compared with sistematic biopsies prostate cancer detection. Initial experience]. Arch Esp Urol. 2020;73(3):192-201.	Wrong study design
Borghesi M, Bianchi L, Barbaresi U, Vagnoni V, Corcioni B, Gaudiano C, et al. Diagnostic performance of MRI/TRUS fusion-guided biopsies vs. systematic prostate biopsies in biopsy-naive, previous negative biopsy patients and men undergoing active surveillance. Minerva Urol Nephrol. 2021;73(3):357-66.	Wrong intervention: out of scope SF
Borkowetz A, Hadaschik B, Platzek I, Toma M, Torsev G, Renner T, et al. Prospective comparison of transperineal magnetic resonance imaging/ultrasonography fusion biopsy and transrectal SB in biopsy-naive patients. BJU Int. 2018;121(1):53-60.	SF vs SB, insufficient data for inclusion in indirect comparison.
Borkowetz A, Platzek I, Toma M, Laniado M, Baretton G, Froehner M, et al. Comparison of systematic transrectal biopsy to transperineal magnetic resonance imaging/ultrasound-fusion biopsy for the diagnosis of prostate cancer. BJU Int. 2015;116(6):873-9.	Wrong study design
Borkowetz A, Platzek I, Toma M, Renner T, Herout R, Baunacke M, et al. Evaluation of Prostate Imaging Reporting and Data System classification in the prediction of tumor aggressiveness in targeted magnetic resonance imaging/ultrasound-fusion biopsy. Urol Int. 2017;99(2):177-85.	Wrong study design
Borkowetz A, Renner T, Platzek I, Toma M, Herout R, Baunacke M, et al. Evaluation of transperineal magnetic resonance imaging/ultrasound-fusion biopsy compared to transrectal SB in the prediction of tumour aggressiveness in patients with previously negative biopsy. Urol Int. 2019;102(1):20-6.	Wrong study design
Brock M, Loppenberg B, Roghmann F, Pelzer A, Dickmann M, Becker W, et al. Impact of real-time elastography on magnetic resonance imaging/ultrasound fusion guided biopsy in patients with prior negative prostate biopsies. J Urol. 2015;193(4):1191-7.	Wrong intervention: out of scope SF
Brock M, von Bodman C, Palisaar J, Becker W, Martin-Seidel P, Noldus J. Detecting prostate cancer. A prospective comparison of systematic prostate biopsy with targeted biopsy guided by fused MRI and transrectal ultrasound. Dtsch Arztebl Int. 2015;112(37):605-11.	Wrong intervention: out of scope SF
Brock M, von Bodman C, Palisaar J, Becker W, Martin-Seidel P, Noldus J. Detecting prostate cancer-a prospective comparison of systematic prostate biopsy with targeted biopsy guided by fused MRI and transrectal ultrasound. Dtsch Arztebl Int. 2015;112:605-11.	Wrong intervention: out of scope SF
Brown LC, Ahmed HU, Faria R, El-Shater Bosaily A, Gabe R, Kaplan RS, et al. Multiparametric MRI to improve detection of prostate cancer compared with transrectal ultrasound-guided prostate biopsy alone: the PROMIS study. Health Technol Assess. 2018;22(39):1-176.	Wrong intervention
Bryk DJ, Llukani E, Huang WC, Lepor H. Natural history of pathologically benign cancer suspicious regions on multiparametric magnetic resonance imaging following targeted biopsy. J Urol. 2015;194(5):1234-40.	Wrong study design
Bryk DJ, Llukani E, Taneja SS, Rosenkrantz AB, Huang WC, Lepor H. The role of ipsilateral and contralateral transrectal ultrasound-guided systematic prostate biopsy in men with unilateral magnetic resonance imaging lesion undergoing magnetic resonance imaging-ultrasound fusion-targeted prostate biopsy. Urology. 2017;102:178-82.	Wrong study design
Bukavina L, Tilburt JC, Konety B, Shah ND, Gross CP, Yu JB, et al. Perceptions of prostate MRI and fusion biopsy of radiation oncologists and urologists for patients diagnosed with prostate cancer: results from a national survey. Eur Urol Focus. 2020;6(2):273-9.	Wrong intervention
Califano A, Caputo A, D'Antonio A, Ciccone V, Fabiano M, Maiorino F, et al. The best prostate biopsy sampling system - fusion and SB: a single center experience [published online ahead of print December 29 2021]. Urologia. 2021.	Wrong intervention: out of scope SF
Campa R, Del Monte M, Barchetti G, Pecoraro M, Salvo V, Ceravolo I, et al. Improvement of prostate cancer detection combining a computer-aided diagnostic system with TRUS-MRI targeted biopsy. *Abdom Radiol (NY)* 2019;**44**:264-71.	Wrong comparator

Cash H, Maxeiner A, Stephan C, Fischer T, Durmus T, Holzmann J, et al. The detection of significant prostate cancer is correlated with the Prostate Imaging Reporting and Data System (PI-RADS) in MRI/transrectal ultrasound fusion biopsy [published online ahead of print 21 August 2015]. World J Urol. 2015.	Wrong intervention: out of scope SF
Castellucci R, Linares Quevedo AI, Sanchez Gomez FJ, Cogollos Acuna I, Salmeron Beliz I, Munoz Fernandez de Legaria M, et al. A non-randomized prospective study on the diagnostic performance of perineal prostatic biopsy, directed via diffusion nuclear resonance, in patients with suspected prostate cancer and previous negative transrectal prostate biopsy. Urologia 2021;88:69-76.	Wrong outcome
Castellucci R, Linares Quevedo AI, Sanchez Gomez FJ, Diez Rodriguez J, Cogorno L, Cogollos Acuna I, et al. Prospective nonrandomized study of diagnostic accuracy comparing prostate cancer detection by transrectal ultrasound-guided biopsy to magnetic resonance imaging with subsequent MRI-guided biopsy in biopsy-naive patients. Minerva Urol Nefrol. 2017;69(6):589-95.	CF vs SB, insufficient data for inclusion in indirect comparison.
Cata E, Andras I, Ferro M, Kadula P, Leucuta D, Musi G, et al. Systematic sampling during MRI-US fusion prostate biopsy can overcome errors of targeting-prospective single center experience after 300 cases in first biopsy setting. Transl Androl Uro. 2020;9(6):2510-8.	Wrong study design
Cattarino S, Forte V, Salciccia S, Drudi FM, Cantisani V, Sciarra A, et al. MRI ultrasound fusion biopsy in prostate cancer detection: are randomized clinical trials reproducible in everyday clinical practice? Urologia. 2019;86(1):9-16.	SF vs SB, insufficient data for inclusion in indirect comparison.
Cauni VM, Stanescu D, Tanase F, Mihai B, Persu C. Magnetic resonance/ultrasound fusion targeted biopsy of the prostate can be improved by adding SB. Med Ultrason. 2021;23(3):277-82.	Wrong study design
Celma A, Lopez R, Roche S, Planas J, Regis L, Placer J, et al. Are targeted prostate biopsies ready to replace systematic prostate biopsies? Actas Urol Esp. 2019;43(10):573-8.	Wrong study design
Chang AI, Park BK. New TRUS techniques and imaging features of PI-RADS 4 or 5: influence on tumor targeting. Front Oncol. 2021;11:608409.	CF vs SB, insufficient data for inclusion in indirect comparison
Chang CH, Chiu HC, Lin WC, Ho TL, Chang H, Chang YH, et al. The influence of serum prostate-specific antigen on the accuracy of magnetic resonance imaging targeted biopsy versus saturation biopsy in patients with previous negative biopsy. Biomed Res Int. 2017;2017:7617148.	Wrong study design
Checcucci E, De Cillis S, Amparore D, Garrou D, Aimar R, Piana A, et al. Naive patients with suspicious prostate cancer and positive multiparametric magnetic resonance imaging (mp-MRI): is it time for fusion target biopsy alone? [published online ahead of print June 22, 2021]. J Clin Urol. 2021.	Wrong study design
Checcucci E, Piramide F, Amparore D, De Cillis S, Granato S, Sica M, et al. Beyond the learning curve of prostate MRI/TRUS target fusion biopsy after more than 1000 procedures. Urology. 2021;155:39-45.	Wrong study design
Chen J, Lin Z, Chen J, Lin Q, Chen J, Yan Y. Magnetic resonance imaging-guided transperineal prostate biopsy. J Cancer Res Ther. 2019;15(2):394-7.	Wrong intervention
Chen J, Yi XL, Jiang LX, Wang R, Zhao JG, Li YH, et al. 3-Tesla magnetic resonance imaging improves the prostate cancer detection rate in transrectral ultrasound-guided biopsy. Exp Ther Med. 2015;9(1):207-12.	CF vs SB, insufficient data for inclusion in indirect comparison
Cheng Y, Qi F, Liang L, Zhang L, Cao D, Hua L, et al. Use of prostate systematic and targeted biopsy on the basis of bi-parametric magnetic resonance imaging in biopsy-naive patients. J Invest Surg. 2022;35(1):92-7.	Wrong study design
Chessa F, Schiavina R, Ercolino A, Gaudiano C, Giusti D, Bianchi L, et al. Diagnostic accuracy of the Novel 29 MHz micro-ultrasound "ExactVu™" for the detection of clinically significant prostate cancer: a prospective single institutional study. A step forward in the diagnosis of prostate cancer. Arch Ital Urol Androl. 2021;93(2):132-8.	Wrong population
Claros OR, Tourinho-Barbosa RR, Fregeville A, Gallardo AC, Muttin F, Carneiro A, et al. Comparison of initial experience with transrectal magnetic resonance imaging cognitive guided micro-ultrasound biopsies versus established transperineal robotic ultrasound magnetic resonance imaging fusion biopsies for prostate cancer. *J Urol* 2020;**203**:918-25.	Wrong comparator: MicroUS is not standard practice
Connor MJ, Eldred-Evans D, van Son M, Hosking-Jervis F, Bertoncelli Tanaka M, Reddy D, et al. A multicenter study of the clinical utility of nontargeted systematic transperineal prostate biopsies in patients undergoing pre-biopsy multiparametric magnetic resonance imaging. J Urol. 2020;204(6):1195-201.	SF and CF combined, no separate data per fusion method
Cool DW, Romagnoli C, Izawa JI, Chin J, Gardi L, Tessier D, et al. Comparison of prostate MRI-3D transrectal ultrasound fusion biopsy for first-time and repeat biopsy patients with previous atypical small acinar proliferation. Can Urol Assoc J. 2016;10(9-10):342-8.	SF vs SB, insufficient data for inclusion in indirect comparison.
Cool DW, Zhang X, Romagnoli C, Izawa JI, Romano WM, Fenster A. Evaluation of MRI-TRUS fusion versus cognitive registration accuracy for MRI-targeted, TRUS-guided prostate biopsy. AJR Am J Roentgenol. 2015;204(1):83-91.	Wrong study design
Cornud F, Lefevre A, Camparo P, Barat M, Dumonceau O, Galiano M, et al. Post-MRI transrectal micro-ultrasonography of transition zone PI-RADS > 2 lesions for biopsy guidance [published online ahead of print April 22 2022]. Eur Radiol. 2022;22.	Wrong study design

Cornud F, Lefevre A, Flam T, Dumonceau O, Galiano M, Soyer P, et al. MRI-directed high-frequency (29MhZ) TRUS-guided biopsies: initial results of a single-center study. *Eur Radiol* 2020;**30**:4838-46.	Wrong comparator: MicroUS is not standard practice
Costa DN, Bloch BN, Yao DF, Sanda MG, Ngo L, Genega EM, et al. Diagnosis of relevant prostate cancer using supplementary cores from magnetic resonance imaging-prompted areas following multiple failed biopsies. Magn Reson Imaging. 2013;31(6):947-52.	Wrong study design
Cricco-Lizza E, Wilcox Vanden Berg RN, Laviana A, Pantuck M, Basourakos SP, Salami SS, et al. Comparative effectiveness and tolerability of transperineal MRI-targeted prostate biopsy under local versus sedation. Urology. 2021;155:33-8.	Wrong study design
D'Agostino D, Mineo Bianchi F, Romagnoli D, Giampaoli M, Corsi P, Del Rosso A, et al. MRI/TRUS FUSION guided biopsy as first approach in ambulatory setting: feasibility and performance of a new fusion device. *Arch Ital Urol Androl* 2020;**91**:211-7.	Wrong comparator
Dal Moro F, Zecchini G, Morlacco A, Gardiman MP, Lacognata CS, Lauro A, et al. Does 1.5 T mpMRI play a definite role in detection of clinically significant prostate cancer? Findings from a prospective study comparing blind 24-core saturation and targeted biopsies with a novel data remodeling model. Aging Clin Exp Res. 2019;31(1):115-23.	CF vs SB, insufficient data for inclusion in indirect comparison
de Gorski A, Roupret M, Peyronnet B, Le Cossec C, Granger B, Comperat E, et al. Accuracy of magnetic resonance imaging/ultrasound fusion targeted biopsies to diagnose clinically significant prostate cancer in enlarged compared to smaller prostates. J Urol. 2015;194(3):669-73.	SF vs SB, insufficient data for inclusion in indirect comparison.
Dekalo S, Matzkin H, Mabjeesh NJ. High cancer detection rate using CF - targeted transperineal prostate biopsies. Int Braz J Urol. 2017;43(4):600-6.	Wrong study design
Del Monte M, Leonardo C, Salvo V, Grompone MD, Pecoraro M, Stanzione A, et al. MRI/US fusion-guided biopsy: performing exclusively targeted biopsies for the early detection of prostate cancer. *Radiol Med* 2018;**123**:227-34.	Wrong comparator
Delongchamps NB, Portalez D, Bruguiere E, Rouviere O, Malavaud B, Mozer P, et al. Are magnetic resonance imaging-transrectal ultrasound guided targeted biopsies noninferior to transrectal ultrasound guided systematic biopsies for the detection of prostate cancer? J Urol. 2016;196(4):1069-75	SF vs SB, insufficient data for inclusion in indirect comparison.
Demirtas A, Sonmez G, Tombul ST, Demirtas T. Comparison of pain levels in fusion prostate biopsy and standard TRUS-guided biopsy. Int Braz J Urol 2020;46:557-62.	Wrong outcome
Demirtas T, Gur A, Golbasi A, Sonmez G, Tombul ST, Demirtas A. A prospective study and single-center experience: effectivity of fusion prostate biopsy in biopsy-naive patients. Cureus. 2021;13(10):e19002.	Wrong intervention: out of scope SF
Demirtas T, Sonmez G, Tombul ST, Demirtas A. A single-center experience: does MRI-guided target prostate biopsy meet expectations? Cureus 2019;11:e6160.	Wrong intervention: out of scope SF
Derigs F, Doryumu S, Tollens F, Norenberg D, Neuberger M, von Hardenberg J, et al. A prospective study on inter-operator variability in semi-robotic software-based MRI/TRUS-fusion targeted prostate biopsies. World J Urol. 2022;40(2):427-33.	SF vs SB, insufficient data for inclusion in indirect comparison.
Dianat SS, Carter HB, Schaeffer EM, Hamper UM, Epstein JI, Macura KJ. Association of quantitative magnetic resonance imaging parameters with histological findings from MRI/ultrasound fusion prostate biopsy. Can J Urol. 2015;22(5):7965-72.	Wrong study design
Ding XF, Luan Y, Zhu LY, Xiao Q, Chen J, Chen HP, et al. The application of intraoperative frozen section examination in multiparametric magnetic resonance imaging/transrectal ultrasound fusion prostate biopsy during a major pandemic. Quant Imaging Med Surg. 2022;12(4):2378-84.	Wrong population
Distler F, Radtke JP, Kesch C, Roethke M, Schlemmer HP, Roth W, et al. [Value of MRI/ultrasound fusion in primary biopsy for the diagnosis of prostate cancer]. Urologe A. 2016;55(2):146-55.	SF vs SB, insufficient data for inclusion in indirect comparison.
Distler FA, Radtke JP, Bonekamp D, Kesch C, Schlemmer HP, Wieczorek K, et al. The value of PSA density in combination with PI-RADS™ for the accuracy of prostate cancer prediction. J Urol. 2017;198(3):575-82.	SF vs SB, insufficient data for inclusion in indirect comparison.
Durmus T, Reichelt U, Huppertz A, Hamm B, Beyersdorff D, Franiel T. MRI-guided biopsy of the prostate: correlation between the cancer detection rate and the number of previous negative TRUS biopsies. Diagn Interv Radiol 2013;19:411-7.	Wrong outcome

Durmus T, Stephan C, Grigoryev M, Diederichs G, Saleh M, Slowinski T, et al. [Detection of prostate cancer by real-time MR/ultrasound fusion-guided biopsy: 3T MRI and state of the art sonography]. *Rofo* 2013;**185**:428-33.	Wrong comparator
Egbers N, Schwenke C, Maxeiner A, Teichgraber U, Franiel T. MRI-guided core needle biopsy of the prostate: acceptance and side effects. Diagn Interv Radiol. 2015;21(3):215-21.	Wrong intervention
Eklund M, Jaderling F, Discacciati A, Bergman M, Annerstedt M, Aly M, et al. MRI-targeted or standard biopsy in prostate cancer screening. N Engl J Med. 2021;385(10):908-20.	Wrong population
El-Achkar A, Abou Heidar N, Labban M, Al-Moussawy M, Moukaddem H, Nasr R, et al. MRI/US fusion transperineal versus transrectral biopsy of prostate cancer: outcomes and complication rates, a tertiary medical center experience in the Middle East. Turk J Urol. 2022;48(2):98-105.	Wrong study design
Eldred-Evans D, Burak P, Connor MJ, Day E, Evans M, Fiorentino F, et al. Population-based prostate cancer screening with magnetic resonance imaging or ultrasonography: the IP1-PROSTAGRAM Study. JAMA Oncol. 2021;7(3):395-402.	SF vs SB, insufficient data for inclusion in indirect comparison (lack of separable data on PCa and csPCa cancer detection rates by SF, SB and combined biopsy approaches).
Engehausen DG, Engelhard K, Schwab SA, Uder M, Wach S, Wullich B, et al. Magnetic resonance image-guided biopsies with a high detection rate of prostate cancer. ScientificWorldJournal. 2012;2012:975971.	Wrong intervention
Engelhard K, Kuhn R, Osten A, Bogner K, Dworak A, Lubke L, et al. Impact of magnetic resonance imaging-guided prostate biopsy in the supine position on the detection of significant prostate cancer in an inhomogeneous patient cohort. Scand J Urol. 2016;50(2):110-5.	Wrong intervention
Febres-Aldana CA, Alghamdi S, Weppelmann TA, Lastarria E, Bhandari A, Omarzai Y, et al. Magnetic resonance imaging-ultrasound fusion-targeted biopsy combined with systematic 12-core ultrasound-guided biopsy improves the detection of clinically significant prostate cancer: are we ready to abandon the systematic approach? Urol Ann. 2020;12(4):366-72.	Wrong study design
Fiard G, Hohn N, Descotes JL, Rambeaud JJ, Troccaz J, Long JA. Targeted MRI-guided prostate biopsies for the detection of prostate cancer: initial clinical experience with real-time 3-dimensional transrectal ultrasound guidance and magnetic resonance/transrectal ultrasound image fusion. Urology. 2013;81(6):1372-8.	Wrong intervention: out of scope SF
Garcia Bennett J, Conejero Olesti A, Hurtado Salom C, Rebenaque E, Parada D, Serrano Alcala E, et al. Usefulness of cognitive targeting in multiparametric MRI-guided biopsy to diagnose the dominant lesion in prostate cancer. Radiologia. 2015;57(5):428-33.	Wrong study design
Garcia Bennett J, Vilanova JC, Guma Padro J, Parada D, Conejero A. Evaluation of MR imaging-targeted biopsies of the prostate in biopsy-naive patients. A single centre study. Diagn Interv Imaging 2017;98:677-84.	Wrong outcome
Gayet MC, van der Aa AA, Beerlage HP, Schrier BP, Gielens M, Heesakkers R, et al. Cancer detection rates of systematic and targeted prostate biopsies after biparametric MRI. Prostate Cancer. 2020;2020:4626781.	Wrong intervention: out of scope SF
Gillis CJ, Southall TM, Wilson R, Anderson M, Young J, Hewitt R, et al. The value of magnetic resonance imaging-ultrasound fusion targeted biopsies for clinical decision making among patients with previously negative transrectal ultrasound biopsy and persistent prostate-specific antigen elevation. Can Urol Assoc J. 2022;16(6):E315-20.	Wrong study design
Giyasov SI, Kilov F, Mukhtarov S, Tukhtamishev MH, Inoyatov UN. [To the issue of improving early diagnosis of localized prostate cancer]. Urologiia 2020:66-72.	Irretrievable
Glybochko PV, Alyaev YG, Amosov AV, Enikeev DV, Chinenov DV, Krupinov GE, et al. [Multi-parametric MRI/US fusion guided biopsy for the diagnosis of prostate cancer. ur experience]. Urologiia 2018:98-104.	Irretrievable
Gomez Gomez E, Valero Rosa J, Carrasco Valiente J, Trivino Tarradas F, Anglada Curado F, Lopez Ruiz D, et al. New approach to guide target prostate biopsy: technique and initial experience. Urology. 2018;121:198-9.	Wrong study design
Gordetsky JB, Nix JW, Rais-Bahrami S. Perineural invasion in prostate cancer is more frequently detected by multiparametric MRI targeted biopsy compared with standard biopsy. Am J Surg Pathol. 2016;40(4):490-4.	Wrong study design

Gorin MA, Meyer AR, Zimmerman M, Harb R, Joice GA, Schwen ZR, et al. Transperineal prostate biopsy with cognitive magnetic resonance imaging/biplanar ultrasound fusion: description of technique and early results. World J Urol. 2020;38(8):1943-9.	Wrong study design
Gronberg H, Eklund M, Picker W, Aly M, Jaderling F, Adolfsson J, et al. Prostate cancer diagnostics using a combination of the Stockholm3 blood test and multiparametric magnetic resonance imaging. Eur Urol. 2018;74(6):722-8.	Wrong study design
Gunzel K, Cash H, Buckendahl J, Konigbauer M, Asbach P, Haas M, et al. The addition of a sagittal image fusion improves the prostate cancer detection in a sensor-based MRI /ultrasound fusion guided targeted biopsy. BMC Urol. 2017;17(1):7.	Wrong intervention: out of scope SF
Gunzel K, Magheli A, Baco E, Cash H, Heinrich S, Neubert H, et al. Infection rate and complications after 621 transperineal MRI-TRUS fusion biopsies in local anesthesia without standard antibiotic prophylaxis. World J Urol. 2021;39(10):3861-6.	Wrong study design
Hadaschik BA, Kuru TH, Tulea C, Rieker P, Popeneciu IV, Simpfendorfer T, et al. A novel stereotactic prostate biopsy system integrating pre-interventional magnetic resonance imaging and live ultrasound fusion. J Urol 2011;186:2214-20.	Wrong outcome
Haffner J, Lemaitre L, Puech P, Haber GP, Leroy X, Jones JS, et al. Role of magnetic resonance imaging before initial biopsy: comparison of magnetic resonance imaging-targeted and SB for significant prostate cancer detection. BJU Int. 2011;108(8 Pt 2):E171-8.	CF vs SB, insufficient data for inclusion in indirect comparison
Hakozaki Y, Matsushima H, Kumagai J, Murata T, Masuda T, Hirai Y, et al. A prospective study of magnetic resonance imaging and ultrasonography (MRI/US)-fusion targeted biopsy and concurrent systematic transperineal biopsy with the average of 18-cores to detect clinically significant prostate cancer. BMC Urol. 2017;17(1):117.	Wrong intervention: out of scope SF
Halstuch D, Baniel J, Lifshitz D, Sela S, Ber Y, Margel D. Characterizing the learning curve of MRI-US fusion prostate biopsies. Prostate Cancer Prostatic Dis 2019;22:546-51.	Wrong outcome
Hambrock T, Somford DM, Hoeks C, Bouwense SA, Huisman H, Yakar D, et al. Magnetic resonance imaging guided prostate biopsy in men with repeat negative biopsies and increased prostate specific antigen. J Urol. 2010;183(2):520-8.	Wrong study design
Hamid S, Donaldson IA, Hu Y, Rodell R, Villarini B, Bonmati E, et al. The SmartTarget biopsy trial: a prospective, within-person randomised, blinded trial comparing the accuracy of visual-registration and magnetic resonance imaging/ultrasound image-fusion targeted biopsies for prostate cancer risk stratification. Eur Urol. 2019;75(5):733-40.	Wrong intervention: SmartTarget SF device is out of scope in this appraisal
Hansen NL, Barrett T, Lloyd T, Warren A, Samel C, Bratt O, et al. Optimising the number of cores for magnetic resonance imaging-guided targeted and systematic transperineal prostate biopsy. BJU Int. 2020;125(2):260-9.	SF vs SB, insufficient data for inclusion in indirect comparison.
Hansen NL, Kesch C, Barrett T, Koo B, Radtke JP, Bonekamp D, et al. Multicentre evaluation of targeted and systematic biopsies using magnetic resonance and ultrasound image-fusion guided transperineal prostate biopsy in patients with a previous negative biopsy. BJU Int. 2017;120(5):631-8.	SF vs SB, insufficient data for inclusion in indirect comparison.
Henning GM, Vetter JM, Sterling JA, Andriole GL, Kim IY, Kim EH. Factors associated with higher prostate biopsy yield: when is software-assisted fusion MRI-targeting necessary? Urol Oncol. 2021;39(4):234.e15-e19.	Wrong study design
Herlemann A, Overland MR, Washington SL, Cowan JE, Westphalen AC, Carroll PR, et al. How often does magnetic resonance imaging detect prostate cancer missed by transrectal ultrasound? Eur Urol Focus. 2021;7(6):1268-73.	SF vs SB, insufficient data for inclusion in indirect comparison.
Hofbauer SL, Luger F, Harland N, Plage H, Reimann M, Hollenbach M, et al. A non-inferiority comparative analysis of micro-ultrasonography and MRI-targeted biopsy in men at risk of prostate cancer. BJU Int 2022;129:648-54.	Wrong comparator: MicroUS is not standard practice
Hoffmann MA, Taymoorian K, Ruf C, Gerhards A, Leyendecker K, Stein T, et al. Diagnostic performance of multiparametric magnetic resonance imaging and fusion targeted biopsy to detect significant prostate cancer. Anticancer Res 2017;37:6871-7.	Wrong outcome
Hoffmann MA, Wieler HJ, Jakobs FM, Taymoorian K, Gerhards A, Miederer M, et al. [Diagnostic significance of multiparametric MRI combined with US-fusion guided biopsy of the prostate in patients with increased PSA levels and negative standard biopsy results to detect significant prostate cancer - correlation with the Gleason score]. Nuklearmedizin 2017;56:147-55.	Wrong outcome
Hsieh PF, Chang TY, Lin WC, Chang H, Chang CH, Huang CP, et al. A comparative study of transperineal software-assisted magnetic resonance/ultrasound fusion biopsy and transrectal CF biopsy of the prostate. BMC Urol. 2022;22(1):72.	Wrong study design
Hwang SI, Ahn H, Lee HJ, Hong SK, Byun SS, Lee S, et al. Comparison of accuracies between real-time nonrigid and rigid registration in the MRI-US fusion biopsy of the prostate. Diagnostics (Basel). 2021;11(8):1481.	Wrong intervention: out of scope SF

Hwang SI, Lee HJ, Lee SE, Hong SK, Byun SS, Lee SC, et al. Value of MR-US fusion in guidance of repeated prostate biopsy in men with PSA<10ng/mL. Clin Imaging. 2019;53:1-5.	Wrong study design
Immerzeel J, Israel B, Bomers J, Schoots IG, van Basten JP, Kurth KH, et al. Multiparametric Magnetic Resonance Imaging for the Detection of Clinically Significant Prostate Cancer: What Urologists Need to Know. Part 4: Transperineal Magnetic Resonance-Ultrasound Fusion Guided Biopsy Using Local Anesthesia. Eur Urol 2022;81:110-7.	Wrong outcome
Jacewicz M, Rud E, Galtung KF, Noor D, Baco E. Cancer detection rates in targeted transperineal MRI-TRUS elastic fusion-guided prostate biopsies performed under local anesthesia. Anticancer Res 2021;41:4395-400.	Wrong outcome
Jambor I, Bostrom PJ, Taimen P, Syvanen K, Kahkonen E, Kallajoki M, et al. Novel biparametric MRI and targeted biopsy improves risk stratification in men with a clinical suspicion of prostate cancer (IMPROD Trial). J Magn Reson Imaging 2017;46:1089-95.	Wrong outcome
Jelidi A, Ohana M, Labani A, Alemann G, Lang H, Roy C. Prostate cancer diagnosis: efficacy of a simple electromagnetic MRI-TRUS fusion method to target biopsies. Eur J Radiol. 2017;86:127-34.	Wrong intervention: out of scope SF
Junker D, Schafer G, Heidegger I, Bektic J, Ladurner M, Jaschke W, et al. Multiparametric magnetic resonance imaging/transrectal ultrasound fusion targeted biopsy of the prostate: preliminary results of a prospective single-centre study. Urol Int. 2015;94(3):313-8.	Wrong intervention: out of scope SF
Kadoury S, Yan PK, Xu S, Glossop N, Choyke P, Turkbey B, et al. Realtime TRUS/SF targeted-biopsy for prostate cancer: a clinical demonstration of increased positive biopsy rates. Proceedings of the International Workshop on Prostate Cancer Imaging, September 2010; Sep 24; Beijing, China. bib checked 14102022: Springer-Verlag Berlin; 2010. p. 52-62.	Wrong intervention: out of scope SF
Kam J, Yuminaga Y, Kim R, Aluwihare K, Macneil F, Ouyang R, et al. Does magnetic resonance imaging-guided biopsy improve prostate cancer detection? A comparison of systematic, CF and ultrasound fusion prostate biopsy. Prostate Int. 2018;6(3):88-93.	Wrong study design: retrospective, prospective evidence identified for SF technology
Kasivisvanathan V, Dufour R, Moore CM, Ahmed HU, Abd-Alazeez M, Charman SC, et al. Transperineal magnetic resonance image targeted prostate biopsy versus transperineal template prostate biopsy in the detection of clinically significant prostate cancer. J Urol. 2013;189(3):860-6.	Wrong study design
Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. N Engl J Med. 2018;378(19):1767-77.	SF and CF combined, no separate data per fusion method
Kaufmann B, Saba K, Schmidli TS, Stutz S, Bissig L, Britschgi AJ, et al. Prostate cancer detection rate in men undergoing transperineal template-guided saturation and targeted prostate biopsy. Prostate 2022;82:388-96.	Wrong outcome
Kaufmann S, Kruck S, Kramer U, Gatidis S, Stenzl A, Roethke M, et al. Direct comparison of targeted MRI-guided biopsy with systematic transrectal ultrasound-guided biopsy in patients with previous negative prostate biopsies. Urol Int. 2015;94(3):319-25.	Wrong study design
Kaufmann S, Mischinger J, Amend B, Rausch S, Adam M, Scharpf M, et al. First report of robot-assisted transperineal fusion versus off-target biopsy in patients undergoing repeat prostate biopsy. World J Urol. 2017;35(7):1023-9.	SF vs SB, insufficient data for inclusion in indirect comparison.
Kaushal R, Das CJ, Singh P, Dogra PN, Kumar R. Multiparametric magnetic resonance imaging-transrectal ultrasound fusion biopsies increase the rate of cancer detection in populations with a low incidence of prostate cancer. Investig Clin Urol. 2019;60(3):156-61.	SF vs SB, insufficient data for inclusion in indirect comparison.
Khoo CC, Eldred-Evans D, Peters M, van Son M, van Rossum PSN, Connor MJ, et al. A comparison of prostate cancer detection between visual estimation (cognitive registration) and image fusion (software registration) targeted transperineal prostate biopsy. J Urol. 2021;205(4):1075-81.	Wrong study design: retrospective, and prospective evidence identified for SF technology
Kim YJ, Huh JS, Park KK. Effectiveness of bi-parametric MR/US fusion biopsy for detecting clinically significant prostate cancer in prostate biopsy naive men. Yonsei Med J. 2019;60(4):346-51.	Wrong study design
Klotz L, Chin J, Black PC, Finelli A, Anidjar M, Bladou F, et al. Comparison of multiparametric magnetic resonance imaging-targeted biopsy with systematic transrectal ultrasonography biopsy for biopsy-naive men at risk for prostate cancer: a phase 3 randomized clinical trial. JAMA Oncol. 2021;7(4):534-42.	SF vs SB, insufficient data for inclusion in indirect comparison; MircoUS is not standard practice.
Klotz L, Chin J, Black PC, Finelli A, Anidjar M, Bladou F, et al. Correction to: Comparison of multiparametric magnetic resonance imaging-targeted biopsy with systematic transrectal ultrasonography biopsy for biopsy-naive menat risk for prostate cancer: a phase 3 randomized clinical trial (vol 7, pg 534, 2021). JAMA Oncol. 2021;7(4):639.	Wrong study design
Klotz LH, Haider MS, Chin JI. Correction to: Comparison of multiparametric magnetic resonance imaging-targeted biopsy with systematic transrectal ultrasonography biopsy for biopsy-naive menat risk for prostate cancer: a phase 3 randomized clinical trial (vol 7, pg 534, 2021). JAMA Oncol. 2021;7(7):1074.	Wrong study design

Koelis. Clinical evaluation report. TRINITY & Accessories. ID: 3701-02-07-02-02-V1.0. Lyon: Cadeucum; 2021.	Wrong outcome
Kroenig M, Schaal K, Benndorf M, Soschynski M, Lenz P, Krauss T, et al. Diagnostic accuracy of robot-guided, software based transperineal MRI/TRUS fusion biopsy of the prostate in a high risk population of previously biopsy negative men. Biomed Res Int. 2016;2016:2384894.	Wrong study design
Kuru TH, Roethke MC, Seidenader J, Simpfendorfer T, Boxler S, Alammar K, et al. Critical evaluation of magnetic resonance imaging targeted, transrectal ultrasound guided transperineal fusion biopsy for detection of prostate cancer. J Urol. 2013;190(4):1380-6.	SF vs SB, insufficient data for inclusion in indirect comparison.
Lacetera V, Cervelli B, Cicetti A, Gabrielloni G, Montesi M, Morcellini R, et al. MRI/US fusion prostate biopsy: our initial experience. Arch Ital Urol Androl. 2016;88(4):296-9.	SF vs SB, insufficient data for inclusion in indirect comparison.
Lantz A, Haug ES, Picker W, Crippa A, Jaderling F, Mortezavi A, et al. Effect of information on prostate biopsy history on biopsy outcomes in the era of MRI-targeted biopsies. World J Urol. 2021;39(4):1153-9.	SF vs SB, insufficient data for inclusion in indirect comparison.
Lashey A, Sehri SY, Meibodi AA, Ghafari M, Gholivandan J, Solymani M, et al. Level of prostate cancer diagnosis using MRI-TRUS fusion biopsy in patients with a negative history of trus biopsy. *Nephrourol Mon* 2019;**11**:e93596.	Wrong comparator
Lee AY, Chen K, Law YM, Ho HS, Cheng CW, Yuen JS, et al. Robot-assisted magnetic resonance imaging-ultrasound fusion transperineal targeted biopsy. Urology. 2021;155:46.	Wrong study design
Lee DH, Nam JK, Park SW, Lee SS, Han JY, Lee SD, et al. Visually estimated MRI targeted prostate biopsy could improve the detection of significant prostate cancer in patients with a PSA level <10 ng/mL. Yonsei Med J. 2016;57(3):565-71.	CF vs SB, insufficient data for inclusion in indirect comparison
Lee DJ, Recabal P, Sjoberg DD, Thong A, Lee JK, Eastham JA, et al. Comparative effectiveness of targeted prostate biopsy using magnetic resonance imaging ultrasound fusion software and visual targeting: a prospective study. J Urol. 2016;196(3):697-702.	Wrong population: majority of patients (65%) were on active surveillance
Lee SH, Chung MS, Kim JH, Oh YT, Rha KH, Chung BH. Magnetic resonance imaging targeted biopsy in men with previously negative prostate biopsy results. J Endourol. 2012;26(7):787-91.	CF vs SB, insufficient data for inclusion in indirect comparison
Lian H, Zhuang J, Wang W, Zhang B, Shi J, Li D, et al. Assessment of free-hand transperineal targeted prostate biopsy using multiparametric magnetic resonance imaging-transrectal ultrasound fusion in Chinese men with prior negative biopsy and elevated prostate-specific antigen. BMC Urol. 2017;17(1):52.	Wrong intervention: out of scope SF
Lim LY, Tan GH, Zainuddin ZM, Fam XI, Goh EH, Syaris OS, et al. Prospective evaluation of using multiparametric magnetic resonance imaging in CF prostate biopsy compared to the standard systematic 12-core biopsy in the detection of prostate cancer. Urol Ann. 2020;12(3):276-82.	Wrong population
Liu H, Ruan M, Wang H, Wang H, Li X, Song G. Can fewer transperineal SB cores have the same prostate cancer detection rate as of magnetic resonance imaging/ultrasound fusion biopsy? Prostate Cancer Prostatic Dis. 2020;23(4):589-95.	SF and CF combined, no separate data per fusion method
Luger F, Gusenleitner A, Kaar J, Mayr C, Loidl W. Does 29Mhz micro-ultrasound provide uniform diagnostic accuracy within and beyond the peripheral zone? Annals of Urology and Nephrology 2019;1.	Wrong comparator: MicroUS is not standard practice
Lughezzani G, Maffei D, Paciotti M, Pereirae JG, Sanchez A, Staerman F, et al. Comparison of micro-ultrasound and multiparametric MRI imaging for prostate cancer: A Multicentre prospective analysis. J Urol 2020;203(Supplement 4):e806-e7.	Wrong comparator: MicroUS is not standard practice
Ma Q, Yang DR, Xue BX, Wang C, Chen HB, Dong Y, et al. Transrectal real-time tissue elastography targeted biopsy coupled with peak strain index improves the detection of clinically important prostate cancer. Oncol Lett. 2017;14(1):210-6.	Wrong intervention
Mannaerts CK, Engelbrecht MRW, Postema AW, van Kollenburg RAA, Hoeks CMA, Savci-Heijink CD, et al. Detection of clinically significant prostate cancer in biopsy-naive men: direct comparison of SB, multiparametric MRI- and contrast-ultrasound-dispersion imaging-targeted biopsy. BJU Int. 2020;126(4):481-93.	SF vs SB, insufficient data for inclusion in indirect comparison.
Mannaerts CK, Kajtazovic A, Lodeizen OAP, Gayet M, Engelbrecht MRW, Jager GJ, et al. The added value of SB in men with suspicion of prostate cancer undergoing multiparametric MRI-targeted biopsy. Urol Oncol. 2019;37(5):298.e1-.e9.	Wrong study design
Mariotti GC, Falsarella PM, Garcia RG, Queiroz MRG, Lemos GC, Baroni RH. Incremental diagnostic value of targeted biopsy using mpMRI-TRUS fusion versus 14-fragments prostatic biopsy: a prospective controlled study. Eur Radiol. 2018;28(1):11-6.	Wrong intervention: out of scope SF
Marra G, Zhuang J, Beltrami M, Calleris G, Zhao X, Marquis A, et al. Transperineal freehand multiparametric SF targeted biopsies under local anaesthesia for prostate cancer diagnosis: a multicentre prospective study of 1014 cases. BJU Int. 2021;127(1):122-30.	Wrong intervention: out of scope SF
Marra G, Zhuang J, Marquis A, Zhao X, Calleris G, Kan Y, et al. Pain in men undergoing transperineal free-hand multiparametric magnetic resonance imaging fusion targeted biopsies under local anesthesia: outcomes and predictors from a multicenter study of 1,008 patients. J Urol 2020;204:1209-15.	Wrong outcome

Martel P, Rakauskas A, Dagher J, La Rosa S, Meuwly JY, Roth B, et al. WITHDRAWN: The benefit of adopting Microultrasound in the prostate cancer imaging pathway: a lesion-by-lesion analysis [published online ahead of print March 12 2022]. Prog Urol. 2022	Wrong study design
Maxeiner A, Nest AM, Stephan C, Cash H, Baur ADJ, Fischer T, et al. Additive value of transrectal systematic ventral biopsies in combination with magnet resonance imaging/ultrasound fusion-guided biopsy in patients with 3 or more negative prostate biopsies. Urol Int. 2020;104(3-4):205-13.	Wrong intervention: out of scope SF
Maxeiner A, Stephan C, Fischer T, Durmus T, Kilic E, Asbach P, et al. [Real-time MRI/US fusion-guided biopsy in biopsy-naive and pre-biopsied patients with suspicion for prostate cancer]. Aktuelle Urol. 2015;46(1):34-8.	Wrong intervention: out of scope SF
Mehmood S, Alothman KI, Alwehaibi A, Alhashim SM. Diagnostic efficacy and safety of transperineal prostate targeted and SB: the preliminary experience of first 100 cases. Arch Ital Urol Androl. 2021;93(2):127-31.	SF vs SB, insufficient data for inclusion in indirect comparison.
Mertan FV, Greer MD, Shih JH, George AK, Kongnyuy M, Muthigi A, et al. Prospective evaluation of the Prostate Imaging Reporting and Data System Version 2 for prostate cancer detection. J Urol. 2016;196(3):690-6.	SF vs SB, insufficient data for inclusion in indirect comparison.
Miah S, Hosking-Jervis F, Connor MJ, Eldred-Evans D, Shah TT, Arya M, et al. A Multicentre Analysis of the Detection of Clinically Significant Prostate Cancer Following Transperineal Image-fusion Targeted and Nontargeted Systematic Prostate Biopsy in Men at Risk. Eur Urol Oncol. 2020;3(3):262-9.	Wrong study design
Miah S, Servian P, Patel A, Lovegrove C, Skelton L, Shah TT, et al. A prospective analysis of robotic targeted MRI-US fusion prostate biopsy using the centroid targeting approach. J Robot Surg 2020;14:69-74.	Wrong outcome
Mischinger J, Kaufmann S, Russo GI, Harland N, Rausch S, Amend B, et al. Targeted vs systematic robot-assisted transperineal magnetic resonance imaging-transrectal ultrasonography fusion prostate biopsy. BJU Int. 2018;121(5):791-8.	Wrong study design
Mozer P, Roupret M, Le Cossec C, Granger B, Comperat E, de Gorski A, et al. First round of targeted biopsies using magnetic resonance imaging/ultrasonography fusion compared with conventional transrectal ultrasonography-guided biopsies for the diagnosis of localised prostate cancer. BJU Int. 2015;115(1):50-7.	Wrong study design
Nakanishi Y, Ito M, Fukushima H, Yokoyama M, Kataoka M, Ikuta S, et al. Who can avoid SB without missing clinically significant prostate cancer in men who undergo magnetic resonance imaging-targeted biopsy? Clin Genitourin Cancer. 2019;17(3):e664-71.	SF and CF combined, no separate data per fusion method
Natarajan S, Marks LS, Margolis DJ, Huang J, Macairan ML, Lieu P, et al. Clinical application of a 3D ultrasound-guided prostate biopsy system. Urol Oncol 2011;29:334-42.	Wrong outcome
Novaes MAS, Mota A, Athanazio DA. Real life data of MRI-targeted biopsy - experience from a single nonacademic centre using CF and 1.5 tesla scanning. Scand J Urol. 2020;54(5):387-92.	Wrong study design
Oderda M, Faletti R, Battisti G, Dalmasso E, Falcone M, Marra G, et al. Prostate cancer detection rate with Koelis fusion biopsies versus cognitive biopsies: a comparative study. Urol Int. 2016;97(2):230-7.	Wrong study design: retrospective, and prospective evidence identified for SF technology
Peltier A, Aoun F, Lemort M, Kwizera F, Paesmans M, Van Velthoven R. MRI-targeted biopsies versus systematic transrectal ultrasound guided biopsies for the diagnosis of localized prostate cancer in biopsy naive men. Biomed Res Int. 2015;2015:571708.	Wrong population
Pepe P, Garufi A, Priolo G, Pennisi M. Transperineal versus transrectal MRI/TRUS fusion targeted biopsy: detection rate of clinically significant prostate cancer. Clin Genitourin Cancer. 2017;15(1):e33-6.	Wrong intervention: out of scope SF
Pepe P, Garufi A, Priolo GD, Pennisi M. Multiparametric MRI/TRUS fusion prostate biopsy: advantages of a transperineal approach. Anticancer Res. 2017;37(6):3291-4.	Wrong intervention: out of scope SF
Pinto PA, Chung PH, Rastinehad AR, Baccala AA, Kruecker J, Benjamin CJ, et al. Magnetic resonance imaging/ultrasound fusion guided prostate biopsy improves cancer detection following transrectal ultrasound biopsy and correlates with multiparametric magnetic resonance imaging. J Urol. 2011;186(4):1281-5.	Wrong intervention: out of scope SF
Ploussard G, Aronson S, Pelsser V, Levental M, Anidjar M, Bladou F. Impact of the type of ultrasound probe on prostate cancer detection rate and characterization in patients undergoing MRI-targeted prostate biopsies using CF. World J Urol. 2014;32(4):977-83.	Wrong study design
Pokorny MR, de Rooij M, Duncan E, Schroder FH, Parkinson R, Barentsz JO, et al. Prospective study of diagnostic accuracy comparing prostate cancer detection by transrectal ultrasound-guided biopsy versus magnetic resonance (MR) imaging with subsequent MR-guided biopsy in men without previous prostate biopsies. Eur Urol. 2014;66(1):22-9.	CF vs SB, insufficient data for inclusion in indirect comparison

Porreca A, Del Giudice F, Giampaoli M, D'Agostino D, Romagnoli D, Corsi P, et al. Adding SB to magnetic resonance ultrasound fusion targeted biopsy of the prostate in men with previous negative biopsy or enrolled in active surveillance programs: a prospective single center, randomized study. Medicine. 2020;99(37):e22059.	SF vs SB, insufficient data for inclusion in indirect comparison.
Postema AW, Scheltema MJ, Mannaerts CK, Van Sloun RJ, Idzenga T, Mischi M, et al. The prostate cancer detection rates of CEUS-targeted versus MRI-targeted versus systematic TRUS-guided biopsies in biopsy-naive men: a prospective, comparative clinical trial using the same patients. BMC Urol. 2017;17(1):27.	Wrong study design
Puech P, Rouviere O, Renard-Penna R, Villers A, Devos P, Colombel M, et al. Prostate cancer diagnosis: multiparametric MR-targeted biopsy with cognitive and transrectal US-MR fusion guidance versus SB-prospective multicenter study. Radiology. 2013;268(2):461-9.	Wrong intervention: out of scope SF
Qu HW, Liu H, Cui ZL, Jin XB, Zhao Y, Wang MW, et al. [Focusing on MRI-suspected lesions in targeted transrectal prostate biopsy guided by MRI-TRUS fusion imaging for the diagnosis of prostate cancer]. Zhong Hua Nan Ke Xue. 2016;22(9):782-6.	Wrong intervention
Radtke JP, Boxler S, Kuru TH, Wolf MB, Alt CD, Popeneciu IV, et al. Improved detection of anterior fibromuscular stroma and transition zone prostate cancer using biparametric and multiparametric MRI with MRI-targeted biopsy and MRI-US fusion guidance. Prostate Cancer Prostatic Dis. 2015;18(3):288-96.	Wrong study design
Radtke JP, Kuru TH, Boxler S, Alt CD, Popeneciu IV, Huettenbrink C, et al. Comparative analysis of transperineal template saturation prostate biopsy versus magnetic resonance imaging targeted biopsy with magnetic resonance imaging-ultrasound fusion guidance. J Urol. 2015;193(1):87-94.	SF vs SB, insufficient data for inclusion in indirect comparison.
Rastinehad AR, Abboud SF, George AK, Frye TP, Ho R, Chelluri R, et al. Reproducibility of multiparametric magnetic resonance imaging and fusion guided prostate biopsy: multi-institutional external validation by a propensity score matched cohort. J Urol. 2016;195(6):1737-43.	Wrong study design
Rastinehad AR, Turkbey B, Salami SS, Yaskiv O, George AK, Fakhoury M, et al. Improving detection of clinically significant prostate cancer: magnetic resonance imaging/transrectal ultrasound fusion guided prostate biopsy. J Urol. 2014;191(6):1749-54.	SF vs SB, insufficient data for inclusion in indirect comparison.
Rodriguez Socarras ME, Gomez Rivas J, Cuadros Rivera V, Reinoso Elbers J, Llanes Gonzalez L, Michel Mercado I, et al. Prostate mapping for cancer diagnosis: the Madrid protocol. transperineal prostate biopsies using multiparametric magnetic resonance imaging fusion and micro-ultrasound guided biopsies. *J Urol* 2020;**204**:726-33.	Wrong comparator: MicroUS is not standard practice
Rouviere O, Puech P, Renard-Penna R, Claudon M, Roy C, Mege-Lechevallier F, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. Lancet Oncol. 2019;20(1):100-9.	SF and CF combined, no separate data per fusion method
Salami SS, Ben-Levi E, Yaskiv O, Ryniker L, Turkbey B, Kavoussi LR, et al. In patients with a previous negative prostate biopsy and a suspicious lesion on magnetic resonance imaging, is a 12-core biopsy still necessary in addition to a targeted biopsy? BJU Int. 2015;115(4):562-70.	SF vs SB, insufficient data for inclusion in indirect comparison.
Sarkar D, Nandi D, Gangoli S, Hicks J, Carter P. The decision of targeted, systematic or combined biopsy in a biopsy naive patient for the diagnosis of prostate cancer, can be made on the basis of multiparametric magnetic resonance imaging. J Clin Urol. 2020;13(3):198-204.	CF vs SB, insufficient data for inclusion in indirect comparison
Sathianathen NJ, Warlick CA, Weight CJ, Ordonez MA, Spilseth B, Metzger GJ, et al. A clinical prediction tool to determine the need for concurrent systematic sampling at the time of magnetic resonance imaging-guided biopsy. BJU Int. 2019;123(4):612-7.	Wrong study design
Schlenker B, Apfelbeck M, Armbruster M, Chaloupka M, Stief CG, Clevert DA. Comparison of PIRADS 3 lesions with histopathological findings after MRI-fusion targeted biopsy of the prostate in a real world-setting. Clin Hemorheol Microcirc. 2019;71(2):165-70.	Wrong study design
Schlenker B, Apfelbeck M, Buchner A, Stief C, Clevert DA. MRI-TRUS fusion biopsy of the prostate: quality of image fusion in a clinical setting. Clin Hemorheol Microcirc 2018;70:433-40.	Wrong outcome
Shoji S, Hiraiwa S, Endo J, Hashida K, Tomonaga T, Nakano M, et al. Manually controlled targeted prostate biopsy with real-time fusion imaging of multiparametric magnetic resonance imaging and transrectal ultrasound: an early experience. Int J Urol. 2015;22(2):173-8.	SF vs SB, insufficient data for inclusion in indirect comparison.
Shoji S, Hiraiwa S, Ogawa T, Kawakami M, Nakano M, Hashida K, et al. Accuracy of real-time magnetic resonance imaging-transrectal ultrasound fusion image-guided transperineal target biopsy with needle tracking with a mechanical position-encoded stepper in detecting significant prostate cancer in biopsy-naive men. Int J Urol. 2017;24(4):288-94.	SF vs SB, insufficient data for inclusion in indirect comparison.

Siddiqui MM, George AK, Rubin R, Rais-Bahrami S, Parnes HL, Merino MJ, et al. Efficiency of prostate cancer diagnosis by MR/ultrasound fusion-guided biopsy vs standard extended-sextant biopsy for MR-visible lesions. J Natl Cancer Inst. 2016;108(9):djw039.	SF vs SB, insufficient data for inclusion in indirect comparison.
Siddiqui MM, Rais-Bahrami S, Truong H, Stamatakis L, Vourganti S, Nix J, et al. Magnetic resonance imaging/ultrasound-fusion biopsy significantly upgrades prostate cancer versus systematic 12-core transrectal ultrasound biopsy. Eur Urol. 2013;64(5):713-9.	SF vs SB, insufficient data for inclusion in indirect comparison.
Siddiqui MM, Rais-Bahrami S, Turkbey B, George AK, Rothwax J, Shakir N, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA. 2015;313(4):390-7.	SF vs SB, insufficient data for inclusion in indirect comparison.
Simmons LAM, Kanthabalan A, Arya M, Briggs T, Barratt D, Charman SC, et al. Accuracy of transperineal targeted prostate biopsies, visual estimation and image fusion in men needing repeat biopsy in the PICTURE trial. J Urol. 2018;200(6):1227-34.	Wrong intervention: SmartTarget SF device is out of scope in this appraisal
Socarras MER, Rivas JG, Cuadros V, Elbers JR, Llanes L, Mercado IM, et al. Prostate mapping for cancer diagnosis: the Madrid protocol. Transperineal prostate biopsies combining micro-ultrasound and MPMRI fusion biopsy. J Urol 2020;203:e999-.	Wrong comparator: MicroUS is not standard practice
Song G, Ruan M, Wang H, Fan Y, He Q, Lin Z, et al. How many targeted biopsy cores are needed for clinically significant prostate cancer detection during transperineal magnetic resonance imaging ultrasound fusion biopsy? J Urol 2020;204:1202-8.	Wrong outcome
Sonmez G, Demirtas T, Tombul ST, Akgun H, Demirtas A. Diagnostic efficiency of systemic immune-inflammation index in fusion prostate biopsy. Actas Urol Esp 2021;45:359-65.	Wrong outcome
Sonmez G, Tombul ST, Imamoglu H, Akgun H, Demirtas A, Tatlisen A. Multiparametric SF-guided prostate biopsy in biopsy naive patients: preliminary results from 80 patients. Turk J Urol. 2019;45(3):196-201.	Wrong intervention: out of scope SF
Sonn GA, Chang E, Natarajan S, Margolis DJ, Macairan M, Lieu P, et al. Value of targeted prostate biopsy using magnetic resonance-ultrasound fusion in men with prior negative biopsy and elevated prostate-specific antigen. Eur Urol. 2014;65(4):809-15.	SF vs SB, insufficient data for inclusion in indirect comparison.
Tae JH, Shim JS, Jin HJ, Yoon SG, No TI, Kim JY, et al. Initial experience of magnetic resonance imaging/ultrasonography fusion transperineal biopsy: biopsy techniques and results for 75 patients. Investig Clin Urol. 2018;59(6):363-70.	SF vs SB, insufficient data for inclusion in indirect comparison.
Tomaskovic I, Pezelj I, Bolanca Culo K, Novosel L, Nikles S, Tomic M, et al. Diagnostic value of cognitive-registration multiparametric magnetic resonance guided biopsy for the detection of prostate cancer after initial negative biopsy. Acta Clin Croat 2018;57:40-5.	Wrong outcome
Tonttila PP, Lantto J, Paakko E, Piippo U, Kauppila S, Lammentausta E, et al. Prebiopsy multiparametric magnetic resonance imaging for prostate cancer diagnosis in biopsy-naive men with suspected prostate cancer based on elevated prostate-specific antigen values: results from a randomized prospective blinded controlled trial. Eur Urol. 2016;69(3):419-25.	CF vs SB, insufficient data for inclusion in indirect comparison
Turkay R, Inci E, Yildiz O, Ozgur E, Tasci AI. Cognitive versus magnetic resonance-ultrasound fusion prostate biopsy: which one is worthier to perform? Ultrasound Q. 2020;36(4):345-9.	Wrong intervention: out of scope SF
Valerio M, McCartan N, Freeman A, Punwani S, Emberton M, Ahmed HU. Visually directed vs. software-based targeted biopsy compared to transperineal template mapping biopsy in the detection of clinically significant prostate cancer. Urol Oncol. 2015;33(10):424.e9-16.	Wrong population: only 20% of population are eligible (biopsy naïve or prior negative biopsy). No separable data.
Vezelis A, Platkevicius G, Kincius M, Gumbys L, Naruseviciute I, Briediene R, et al. Systematic and MRI-cognitive targeted transperineal prostate biopsy accuracy in detecting clinically significant prostate cancer after previous negative biopsy and persisting suspicion of malignancy. Medicina 2021;57:57.	Wrong outcome
Westhoff N, Haumann H, Kriegmair MC, von Hardenberg J, Budjan J, Porubsky S, et al. Association of training level and outcome of software-based image fusion-guided targeted prostate biopsies. World J Urol. 2019;37(10):2119-27.	Wrong study design
Wetterauer C, Trotsenko P, Matthias MO, Breit C, Keller N, Meyer A, et al. Diagnostic accuracy and clinical implications of robotic assisted MRI-US fusion guided target saturation biopsy of the prostate. *Sci Rep* 2021;**11**:20250.	Wrong comparator
Wiemer L, Hollenbach M, Heckmann R, Kittner B, Plage H, Reimann M, et al. Evolution of targeted prostate biopsy by adding micro-ultrasound to the magnetic resonance imaging pathway. Eur Urol Focus 2021;7:1292-9.	Wrong comparator: MicroUS is not standard practice
Winoker JS, Wajswol E, Falagario U, Maritini A, Moshier E, Voutsinas N, et al. Transperineal versus transrectal targeted biopsy with use of electromagnetically-tracked MR/US fusion guidance platform for the detection of clinically significant prostate cancer. Urology. 2020;146:278-86.	SF vs SB, insufficient data for inclusion in indirect comparison.
Winther MD, Balslev I, Boesen L, Logager V, Noergaard N, Thestrup KD, et al. Magnetic resonance imaging-guided biopsies may improve diagnosis in biopsy-naive men with suspicion of prostate cancer. Dan Med J. 2017;64(5):A5355.	Wrong intervention: out of scope SF

Yamada Y, Shiraishi T, Ueno A, Ueda T, Fujihara A, Naitoh Y, et al. Magnetic resonance imaging-guided targeted prostate biopsy: comparison between computer-software-based fusion versus CF technique in biopsy-naive patients. Int J Urol. 2020;27(1):67-71.	Wrong study design: retrospective, and prospective evidence identified for SF technology
Zalesky M, Stejskal J, Minarik I, Adamcova V, Babjuk M, Zachoval R. Cancer detection rates and inter-examiner variability of MRI/TRUS fusion targeted biopsy and systematic transrectal biopsy. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2020;164(3):314-9.	Wrong intervention: out of scope SF
Zhang J, Zhu A, Sun D, Guo S, Zhang H, Liu S, et al. Is targeted magnetic resonance imaging/transrectal ultrasound fusion prostate biopsy enough for the detection of prostate cancer in patients with PI-RADS >=3: results of a prospective, randomized clinical trial. J Cancer Res Ther. 2020;16(7):1698-702.	Wrong intervention: out of scope SF
Zhang Q, Wang W, Yang R, Zhang G, Zhang B, Li W, et al. Free-hand transperineal targeted prostate biopsy with real-time fusion imaging of multiparametric magnetic resonance imaging and transrectal ultrasound: single-center experience in China. Int Urol Nephrol. 2015;47(5):727-33.	Wrong intervention: out of scope SF
Zhang Q, Wang W, Zhang B, Shi J, Fu Y, Li D, et al. Comparison of free-hand transperineal mpMRI/TRUS fusion-guided biopsy with transperineal 12-core SB for the diagnosis of prostate cancer: a single-center prospective study in China. Int Urol Nephrol. 2017;49(3):439-48.	Wrong intervention: out of scope SF
Zhou Y, Zhou Z, Li Q, Xu Y, Sun H, Xiao Y, et al. Diagnostic accuracy of magnetic resonance-guided prostate biopsy and template-guided transperineal saturation biopsy. Medicine. 2018;97(38):e12495.	SF and CF combined, no separate data per fusion method
Zhu G, Wang Q. Comparisons between magnetic resonance/ultrasound fusion-guided biopsy and standard biopsy in the diagnosis of prostate cancer. A prospective cohort study. Medicine 2018;97:e11962.	Wrong outcome
Zogal P, Sakas G, Rosch W, Baltas D. BiopSee - transperineal stereotactic navigated prostate biopsy. J Contemp Brachytherapy 2011;3:91-5.	Wrong outcome

SF: software fusion; CF: cognitive fusion; SB: systematic biopsy; MicroUS: Micro ultrasound

Table 60 Summary of key studies of FusionVu, Fusion Bx 2.0 and bkFusion excluded from the systematic review of clinical effectiveness

Study	Country	Design	Population	N	MRI Fusion device	Comparator	Findings summary
Cash (2022)[164]	Germany, Austria, N'lands and USA	Retrospective	NR	NR	Fusion Vu (ExactImaging)	N/A	13 practitioners conducted over 1190 micro-ultrasound biopsy procedures during a four-stage training programme. The csPCa detection rate improved from 40% at the start of the training programme to 57% at the end of the training programme (where <50 biopsies were performed before analysis) [p < 0.01]. The final stage of the training programme was independently associated with a higher csPCa detection rate after correcting for overall risk factors (OR 1.95; P = 0.03).
Cornud (2020)[170]	France	Prospective	BN, RB, AS Elevated or rising PSA	118	Fusion Vu (ExactImaging) Systematic biopsy conducted at physicians' discretion	N/A	Fusion biopsy was only used for MRI+/microUS− lesions (n = 13). There were no cases of csPCa, but two cases (15%) of non-significant PCa (Gleason 3+3 and cancer core length ≤ 3mm).
Klotz (2020)[171]	Canada, Italy, Spain, France, German, Austria, USA	Prospective	Prior mpMRI Elevated PSA and/or abnormal DRE	62 (of 1140)	Biojet, Uronav, FusionVu, Hitachi Cognitive fusion Systematic biopsy	N/A	Individuals who had prior mpMRI underwent ExactVu micro-ultrasound-guided biopsy. Biopsies were taken from both mpMRI targets (PI-RADS >3) and micro-ultrasound targets (PRIMUS >3), using either cognitive fusion or software fusion systems. 4/11 centres used software fusion devices, and 62 patients were biopsied using FusionVu. Overall, PCa (Gleason 3+3) was identified in 61% of patients, Gleason ≥ 3+4 was detected in 39% of patients and Gleason ≥ 4+3 was detected in 19% of patients.
Wiemer (2020)[172]	Germany	Prospective	BN, RB Included men at 'clinical risk of prostate cancer'	159	Fusion Vu (ExactImaging) Systematic biopsy	MircoUS-guided biopsy Systematic biopsy	MicroUS positive lesions blinded to the mpMRI results and targeted independently of the mpMRI lesions. The lesion-level positive predictive value for csPCa was 41% for microUS-targeted biopsies and 30% for MRI-targeted biopsies (p = 0.02). MRI targets upgraded the gleason grade compared to systematic biopsy in 34 patients (21%). Micro-US targeting led to an upgrade in

Study	Country	Design	Population	N	MRI Fusion device	Comparator	Findings summary
							gleason grade in 9.4% more patients than MRI targeting (95% CI 2.2–16.5%, p = 0.005). Limited data comparing FusionVu and systematic biopsy (no cross tabulation and different number of patients).
Hofbauer (2021)[173]	Germany	Prospective	BN, RB Included men at 'clinical risk of prostate cancer'	203	Fusion Vu (ExactImaging) Systematic biopsy	MircoUS-guided biopsy Systematic biopsy	The PCa and csPCA detection rate was 63% (127/203 patients) and 39% (79/203). MicroUS-targeted biopsies detected csPCa in 58/79 (73%) patients, compared to MRI-guided biopsies which detected 60/79 (76%). MicroUS was non-inferior compared to mpMRI and detected 97% of csPCa detected by MRI-guided biopsy. (95% CI 80-116%, p = 0.023). MRI-guided biopsy detected 7/79 (9%) csPCa alone, compared to 5/79 (6%) in microUS-guided biopsy. Systematic biopsy alone detected 4/79 (4%) csPCa cases.
Perlis (2020)[174]	Canada	Retrospective	Biopsy experience not reported PI-RADS ≥3, rising or abnormal PSA	51	Fusion Bx 2.0 (Focal Healthcare). Systematic biopsy conducted at physicians' discretion		Early experience study. The PCa (Gleason 3+3) detection rate for PI-RADS 3/4/5 lesions was 22 %/61%/83% respectively. For csPCa (Gleason ≥3+4) detection rate for PI-RADS 3/4/5 lesions was 6%/47%/75%, respectively. No major complications
Miah (2020)[175]	UK	Retrospective	BN, RB Elevated PSA or abnormal DRE	640	bkFusion (BK Medical and MIM) MIM Symphony-DX	Systematic Biopsy	In the total population (n=640), csPCa (Gleason ≥4 + 3 or any grade ≥6 mm) was detected in 41.1% of cases. 357 men who underwent both software fusion and systematic biopsies. The csPCa detection rate by software fusion was 41.1%. csPCa was detected by systematic biopsy alone in three (0.8%) patients.
Immerzeel (2021)[176]	N'lands	Prospective	BN PSA >3ng/mL and/or abnormal DRE, PI-RADS ≥3	1097	BK-Flex Focus 800, BK-3000 (in combination with systematic biopsy)	Perilesional Biopsy	Grade ≥2 (Clavien-Dindo) adverse events were reported in 0.73% patients (8/1097). PCa (Gleason 3+3) and csPCa (Gleason ≥3+4) detected in 84% and 66% of patients respectively. Additional perilesional biopsies were performed in 958/1097 patients, which resulted in the histopathological upgrading in 5.7% of men.

NR, not reported; BN, Biopsy naïve; RB, repeat biopsy; PCa, prostate cancer; csPCa, clinically significant prostate cancer; PI-RADS Prostate Imaging Reporting & Data System; US, ultrasound; DRE, digital rectal examination; PSA, prostate specific antigen; AS, active surveillance.

APPENDIX 4. ADDITIONAL STUDY AND POPULATION CHARACTERISTICS

Table 61 Operator experience in studies included in the systematic review

Study	Operator experience
Cornud (2018)[93]	>10 yrs in MRI & elastic SF
Delongchamps (2013)[98]	'Experienced uroradiologist'
FUTURE (2019)[31, 102]	'Performed by five urologists and expert-trained PhD candidates having at least 6 mo of experience, including 3 mo of experience under expert supervision'
Hansen (2018)[95]	SF (Centre 1): several years' experience of TP biopsy. SF (Centre 2): Supervised Residents. CF: 1/5 urologists
Izadpanahi (2021)[82]	'Experience of performing at least 2000 targeted prostate biopsies'
PAIREDCAP (2019)[88]	'Experienced'
PROFUS (2014)[97]	NR
Stabile (2018)[89]	Urologists had performed at least 200 prostate biopsies but were naïve for TB techniques.
Kaufman (2018)[91, 101]	NR
Liang (2020)[85]	Experienced urologist with more than one year experience
Lockhart (2022)[100]	Experienced radiologist
Monda (2018)[90]	NR
Rabah (2021)[84]	NR
Ferriero (2022)[81]	9 years experience
Sokolakis (2021)[83]	4 operators with no prior experience on mpMRI/TRUS fusion PB, 2 trainees who accomplished 40 TRUS-guided biopsies; and two senior urologists who had done over 250 TRUS-guided biopsies
Alberts (2018)[80]	
Albisinni (2018)[94]	Single operator who performs >100 TBs each year with ?20 years experience
Filson (2016)[96]	NR
Fourcade (2018)[92]	NR
Wajswol (2020)[87]	NR
Gomez-Ortiz (2022)[99]	NR
Kulis (2020)[86]	NR
Thangarasu (2021)[79]	NR

NR: Not reported

APPENDIX 5. QUALITY ASSESSMENT

Table 62 Risk of bias and applicability assessment with rationale

Study	Tests	Refrence std or tests to estimate total positive rates	Risk of bias (QUADAS-C)					Applicability concerns (QUADAS-2)			
			P	I	R	FT	Comments	P	I	R	Comments
Alberts 2018[80]	SF (Koelis Urostation) SB	SF+SB	✓	✓	✗	✓	SF performed after SB within the same examination, by the same operator; no blinding.	✓	?	✓	Equivalence of Urostation (out of scope) with Koelis Trinity (in scope) is uncertain. Anaesthesia method NR.
Albisinni 2018[94]	SF (Koelis Urostation) SB	SF+SB	✓	✓	✗	✓	SF performed after SB within the same examination, by the same operator; no blinding.	✗	?	✓	All patients had one prior negative TRUS. Equivalence of Urostation (out of scope) with Koelis Trinity (in scope) is uncertain. Anaesthesia method NR.
Cornud 2018[93]	SB CF	SF+CF±SB	✓	✗	✓	✗	Although SF and CF were conducted by a separate operator, both were conducted within the same session and tracks from the first method may have been visible. 12 out of 100 patients were not considered for analysis because of missing data (n = 6) or difficulties in extracting the information from the Digital Imaging and Communications in Medicine archives of the biopsy procedure (n = 6).	?	✓ ?	?	47% referred following a prior negative SB. Urostation (TR) is not within scope. Equivalence with Koelis Trinity (in scope) is uncertain. Reference standard informed by both index tests +SB in unknown number of patients
Delongchamps 2013[98]	SF (Urostation Touch, Koelis) CF	SF+SB CF+SB	✗	✓	✗	✓	Consecutive series, unpaired, no matching. Targeted biopsies performed after SB within the same examination, by the same operator; no blinding. Different reference standards were used in relative comparisons (CF+SB vs SF+SB)	✓	? ✓	✓	Applicability of Koelis Urostation to Koelis Trinity is uncertain. Anaesthesia method NR.
Elkhoury 2019[88] (PAIREDCAP)	SF (Artemis) CF	SF+CF+SB	✓	✗	✓	✓	SB, followed by CF, then SF by same operator in the same session. SB operator blinded to MRI report, but no blinding of SF operator to CF tracks.	✓	✓	✓	

Ferriero 2022[81]	SF (Urostation, KOELIS) SF (Biojet)	SF (Urostation, KOELIS) SF (Biojet)	✗	✗	✗	✓	Significant differences in characteristics of two study cohorts (including age, positive DRE and n of target cores), although attempts were made to adjust with propensity score matching (PSM). After adjustment, significant differences remained in median n of target cores (4 (IQR 4-6) for Urostation, vs 6 (4-6) for Biojet. N following PSM reduced from 103 to 83 (Urostation) and 211 to 83 (Biojet). Unclear if anaesthesia and biopsy routes differed between the two index tests. Different reference standard between study arms, only informed by one of two index tests.	✓	?	✗	Applicability of Urostation to Koelis Trinity is unknown. Anaesthesia type unclear. Biopsy positivity rates were not informed by SB, but only by SF biopsies.
Filson 2016[96]	SF (Artemis) SB	SF+SB	✓	✓	✗	✓	SB performed after SF within the same examination, by the same operator; no blinding.	✓	?	✓	Biopsy route and anaesthesia method NR
Fourcade 2018[92]	SF (Koelis Urostation) SB	SF+SB	✓	✓	✗	✓	No blinding; biopsy method order NR.	?	?	✓	Half of the patients had a prior negative biopsy. Biopsy route and anaesthesia method NR. Applicability of Urostation to Koelis Trinity is unknown.
FUTURE[31]	SF (Biopsee) CF	SF CF	?	✗	✗	✓	RCT, no reporting of allocation concealment; higher proportion of posterior lesions in cog (59%) vs software fusion arm (44%). Different routes and anaesthesia methods between arms (TP & GA for SF, vs. TR and LA for CF) No SB; test positivity informed by index test, which by design differed between the two arms.	✗	✗ ✓	✗	Only includes individuals with prior negative SB. SF conducted under GA. Positivity rate was only informed by targeted biopsy (index test).
Gomez-Ortiz 2022[99]	CF SB	CF+SB	✓	✓	✗	✓	SB performed after CF within the same examination, by the same operator; no blinding.	✗	?	✓	All patients had prior negative biopsy. Anaesthesia method NR.
Hansen 2018[95]	SF (Biopsee) CF	SF+SB CF+SB	✗	✓	✗	✗	Allocation to SF or CF according to study centre. Participant allocation not randomized, no matching. Different reference standards used between centres (CF+SB in 1 centre, SF+SB in 2 centres).	✓	✗	✗	All index test and reference standard biopsies performed under GA.

							Significant number of participants in centre III were excluded from the analysis due to process errors.				
Izadpanahi 2021[82]	SF (Artemis)+SB	SF+SB	✓	✓	✗	✓	Different reference standard test between arms.	✓	✓	✓	
	CF+SB	CF+SB									
Kaufmann 2018[101]	SF (iSR'obot Mona Lisa)	SF	✗	✗	✗	✓	Assignment to SF (TP, GA) or CF (TR, LA) based on patient preference, & statistically significant differences between arms in PSA density, median lesion size, and cancer positive rate, though nearest neighbour matching was performed. SF conducted transperineally under GA, CF transrectally under LA. Different reference standards used between study arms (SF+SB or CF+SB)	?	✗	✗	Large proportion of prior negative biopsy patients (40%). Positive DRE excluded. SF conducted under GA. Cancer rate was only informed by targeted biopsy (index test).
	CF	CF							✓		
Kulis 2020[86]	CF	CF+SB	✓	✓	✗	✓	SB performed after CF within the same examination, by the same operator; no blinding.	✗	✓	✓	All patients had prior negative TRUS.
	SB										
Liang 2020[85]	SF (BK)	SF	?	✓	✗	✓	No random allocation; criteria for assignment to SF and CF NR; no significant differences in characteristics between SF and CF arms. No systematic biopsy; cancer rates only informed by targeted biopsy, which by design differed between the study arms (either SF or CF)	✓	✓	✗	Positivity rate was only informed by targeted biopsy (index test).
	CF	CF									
Lockhart 2022[100]	SF (BK/MIM)	SF+SB	✗	✓	✗	✓	Retrospective, criteria for assignment to FS and CF NR; significant differences in characteristics between the two study arms, including mean PSA, AS, median ISUP, mean n of cores per case, CsPCa rates. No blinding; biopsy method order NR. Different reference standard used between arms (SF+SB, vs CF+SB)	✓	?	✓	Biopsy route and anaesthesia method NR.
	CF	CF+SB									

Monda 2018[90]	SF (Uronav) CF	SF+SB CF+SB	✗	✓	✗	✓	Assignment to SF and CF determined by time of introduction of SF to practice. Significant difference in percentage of biopsy naïve (SF: 36%; CF: 27%). Targeted and SB performed in same session, order NR, no blinding reported. Different reference standards between study arms due to design (SF+SB, or CF+SB).	✗	?	✓	Only 36% of SF and 27% of CF were biopsy naïve; 18% and 21% were on AS respectively. Biopsy route and anaesthesia method NR.
PROFUS[97]	SF (Artemis) CF	SF+CF	✓	✗	?	✓	Although CF was blinded to the software fusion targets and conducted by a separate operator, the risk that biopsy tracks from SF biopsy may have influenced the placement of CF cores cannot be excluded. Results for SF+SB and CF+SB, or comparisons between each targeted method with SF+CF+SB NR.	✓	✓	✗	Results for SF+SB and CF+SB, or comparisons between each targeted method with SF+CF+SB NR.
Rabah 2021[84]	SF (Biojet) SF (Artemis)	SF (Biojet) SF (Artemis)	?	✗	✗	✓	Insufficient details on random allocation method and allocation concealment; unclear why a larger number of patients was randomized to TRUSBx (n=165) than TPBx (n=142); no baseline imbalances reported, although no data on lesions location reported. GA was peformed for the TPBx arm only; N of biopsies taken was higher in TRUSBx arm (n=403) compared with TPBx (n=338). Positive rates only informed by one index test in each arm. Each arm had a different software fusion method, route and anaesthesia type.	✓	✗ ✓	✗	All biojet biopsies performed under GA. Positive rates only informed by one index test in each arm. SB (12 core) were conducted for all patients but not included as part of ref std.
Sokolakis 2021[83]	SF (Biojet) SF (Uronav) SF (Koelis Trinity)	SF (Biojet) SF (Uronav) SF (Koelis Trinity)	✗	✓	✗	✓	No randomisation, consecutive series. Small sample size in each arm; no statistically significant differences in reported characteristics, though difference in % with previous biopsy (0 in Trinity arm, vs. 40% in Uronav and 22% in biojet arm. Different test for positive rate estimates for each cohort; SB was not incorporated to the results.	✓	✓	✗	Positivity rate was only informed by targeted biopsy (index test).

Stabile 2018[89]	SF (Biojet)	SF+SB	✗	✗	✗	✓	Unpaired, unmatched design; choice of TB method (including route) at operator's discretion; statistically significant difference in age, PSA, median n of targets per lesion, and previous biopsy between software fusion and cog fusion cohorts (p<0.05). Median target cores per MRI was higher in the software fusion cohort (3, IQR [2-3] than the cognitive biopsy cohort (2 [2-5]) (p<0.001), which may favour the fusion biopsy group. Different reference standards between arms (SB+cog vs SB+software fusion) and no blinding of SB operator.	?	✗	✓	46% prior negative biopsy. All three urologists were naïve to targeted biopsy techniques. Evidence of significant learning curve provided for all targeted biopsy approaches. Anaesthesia method NR.
	CF	CF+SB									
Thangarasu 2021[79]	CF	CF+SB	✓	✓	✗	✓	SB performed after CF within the same examination, by the same operator; no blinding.	✓	✓	✓	
	SB										
Wasjwol 2020[87]	SF (Uronav)	SF+SB	✓	✓	✗	✓	SB performed after SF within the same examination, by the same operator; no blinding.	?	✓	✓	49% had prior negative biopsy.
	SB										

P: patient selection; I : index test; R : reference standard/test(s) used to derive overall biopsy positive rates; FT : flow and timing; SF : software fusion; CF : cognitive fusion; TP: transperineal; TR: transrectal; GA: general anaesthesia; LA: local anaesthesia; NR: not reported; SSB: saturation biopsy

✓ indicates low risk; ✗ indicates high risk; ? indicates unclear risk

APPENDIX 6. ADDITIONAL NETWORK META-ANALYSIS DATA AND RESULTS

Data for additional analyses

Table 63 Data for NMA comparing the number of prostate cancers detected.

	intervention			number of patients			number of cancers		
Study	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**
PAIREDCAP (2019)[88]	CF	SB	Artemis	248	248	248	154	196	177
Izadpanahi (2021)[82]	CF + SB	Artemis + SB	NA	100	99	NA	31	44	NA
Wajswol (2020)[87]	SB	Uronav	Uronav + SB	169	169	169	116	120	133
Thangarasu (2021)[79]	CF	SB	CF + SB	75	75	75	34	40	43
Kulis (2020)[86]	CF	SB	CF + SB	63	63	63	33	30	38
Cornud (2018)[93]	CF	Urostation	NA	88	88	NA	31	40	NA
FUTURE (2019)[31]	CF	BiopSee	NA	78	79	NA	34	39	NA
PROFUS (2014)[97]	CF	Artemis	NA	125	125	NA	40	45	NA
Albisinni (2018)[94]	SB	Urostation	Urostation + SB	74	74	74	33	35	42
Fourcade (2018)[92]	SB	Urostation	Urostation + SB	191	191	191	88	85	106
Gomez-Ortiz (2022)[99]	CF	SB	CF + SB	111	111	111	42	30	46
*Rabah (2021)[84]	Artemis	Biojet	NA	165	142	NA	48	64	NA
Alberts (2018)[80]	SB	Urostation	Urostation + SB	48	48	48	25	28	32
Filson (2016)[96]	SB	Artemis	Artemis + SB	538	538	538	244	228	286

* Study only included in analyses with individual device effects as it compares two software fusion devices;
SF: software fusion; CF: cognitive fusion; SB: systematic biopsy; NA: not applicable

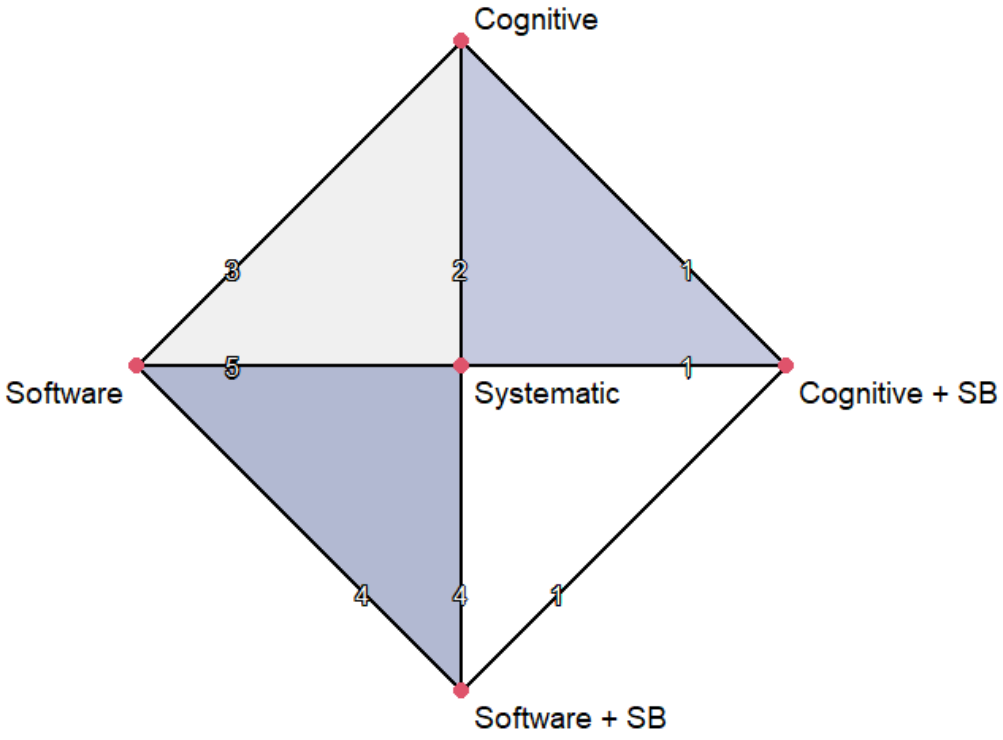
Table 64 Data for NMA comparing the number of clinically significant prostate cancers detected.

	intervention			number of patients			number of clinically significant cancers		
Study	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**	**arm 1**	**arm 2**	**arm 3**

PAIREDCAP (2019)[88]	CF	SB	Artemis	248	248	248	116	150	134
Izadpanahi (2021)[82]	CF + SB	Artemis + SB	NA	100	99	NA	12	19	NA
FUTURE (2019)[31]	CF	BiopSee	NA	78	79	NA	26	27	NA
PROFUS (2014)[97]	CF	Artemis	NA	125	125	NA	24	29	NA
Albisinni (2018)[94]	SB	Urostation	Urostation + SB	74	74	74	21	25	42
Fourcade (2018)[92]	SB	Urostation	Urostation + SB	191	191	191	52	60	106
Gomez-Ortiz (2022)[99]	CF	SB	CF + SB	111	111	111	23	21	46
*Rabah (2021)[84]	Artemis	Biojet	NA	165	142	NA	21	46	NA
Alberts (2018)[80]	SB	Urostation	Urostation + SB	48	48	48	14	17	19
Filson (2016)[96]	SB	Artemis	Artemis + SB	538	538	538	130	160	186

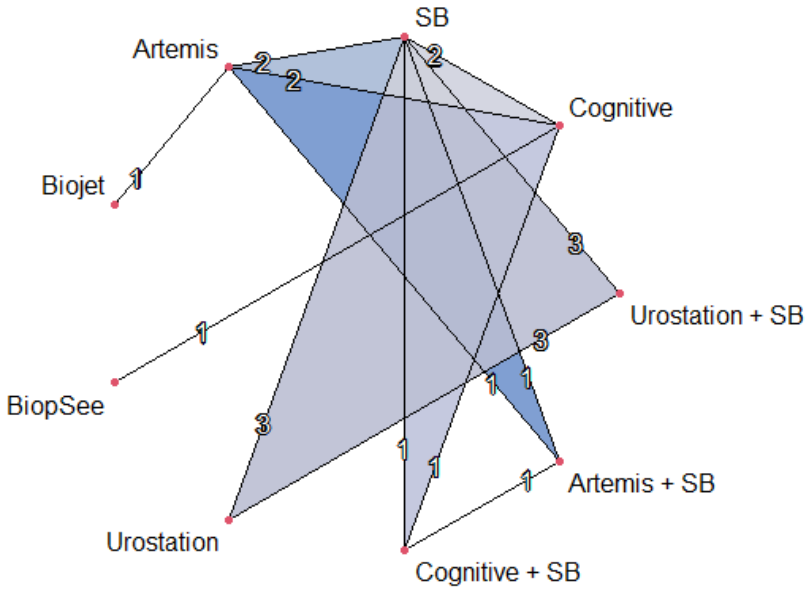
* Study only included in analyses with individual device effects as it compares two software fusion devices;
SF: software fusion; CF: cognitive fusion; SB: systematic biopsy; NA: not applicable

Figure 12 Network of biopsy types and devices compared for clinically significant prostate cancer detection, under the assumption of a common effect for different software fusion devices.



Lines represent comparisons made in studies, numbers on the lines show how many studies included that comparison and shaded areas represent multi-arm studies.
Abbreviations: SB, systematic biopsy.

Figure 13 Network of biopsy types and devices compared for clinically significant prostate cancer detection.



Lines represent comparisons made in studies, numbers on the lines show how many studies included that comparison and shaded areas represent multi-arm studies.
Abbreviations: SB, systematic biopsy.

Results from additional analyses: Tables

Table 65 Probabilities (median and 95%CrI) of being classified at different ISUP grades for biopsy-naïve patients.

	Artemis probabilities from Filson (2016)[96] biopsy-naïve data						Artemis + SB probabilities from Filson (2016)[96] biopsy-naïve data		
ISUP	**Cognitive**		**Systematic**		**Software***		**Cognitive + SB**		**Software + SB***
No cancer	0.36	(0.29, 0.44)	0.25	(0.21, 0.29)	0.29		0.41	(0.21, 0.56)	0.36
1	0.20	(0.15, 0.25)	0.21	(0.17, 0.26)	0.17		0.21	(0.10, 0.33)	0.22
2	0.18	(0.13, 0.25)	0.28	(0.23, 0.33)	0.28		0.10	(0.03, 0.23)	0.22
3	0.15	(0.10, 0.23)	0.15	(0.10, 0.22)	0.16		0.21	(0.06, 0.59)	0.12
4-5	0.10	(0.06, 0.17)	0.10	(0.06, 0.17)	0.10		0.02	(0.00, 0.18)	0.08

* Assumed underlying baseline probabilities.

Table 66 Model 1b: Odds ratios (median and 95% CrI) of being classified as ISUP grades 1 to 4-5 compared to being categorised as having no cancer, for the different (single) biopsy methods, compared to cognitive fusion biopsy; and categorisations using the different biopsy methods combined with systematic biopsy, compared to cognitive fusion plus systematic biopsy.

	Compared to cognitive fusion biopsy										**Compared to cognitive fusion biopsy plus systematic biopsy**			
ISUP	**SB**		**Artemis**		**Biojet**		**BiopSee**		**Urostation**		**Artemis + SB**		**Urostation + SB**	
No cancer	REFERENCE										REFERENCE			
1	1.54	(1.06, 2.24)	1.04	(0.72, 1.52)	1.04	(0.49, 2.24)	1.65	(0.61, 4.73)	1.21	(0.67, 2.15)	1.17	(0.71, 1.94)	1.28	(0.65, 2.53)
2	2.28	(1.50, 3.47)	1.92	(1.26, 2.95)	NE	NE	NE	NE	3.03	(1.04, 8.94)	2.57	(0.94, 8.13)	3.31	(0.78, 15.77)
3	1.41	(0.82, 2.41)	1.31	(0.78, 2.22)	NE	NE	NE	NE	NE	NE	0.65	(0.12, 2.90)	NE	NE
4-5	1.54	(0.83, 2.86)	1.22	(0.65, 2.29)	NE	NE	NE	NE	NE	NE	4.41	(0.46, 150.05)	NE	NE

No results can be obtained for Uronav or Uronav + SB due to lack of detailed ISUP grade reporting.
SB: systematic biopsy; ISUP grade: International Society of Urological Pathology; NE, not estimable (due to data sparseness).

Table 67 Model 1b: Probabilities (median and 95%CrI) of being classified at different ISUP grades for biopsy-naïve patients based on the independent effects analysis.

	Artemis probabilities from Filson (2016)[96] biopsy-naïve data						Artemis + SB probabilities from Filson (2016)[96] biopsy-naïve data		
ISUP	**Cognitive**		**SB**		**Artemis***		**Cognitive+SB**		**Artemis+SB***
No cancer	0.54	(0.46, 0.61)	0.41	(0.35, 0.47)	0.47		0.41	(0.21, 0.56)	0.36
1	0.18	(0.13, 0.24)	0.21	(0.17, 0.26)	0.16		0.21	(0.10, 0.33)	0.22
2	0.12	(0.08, 0.16)	0.20	(0.16, 0.25)	0.20		0.10	(0.03, 0.23)	0.22
3	0.09	(0.06, 0.14)	0.10	(0.06, 0.15)	0.11		0.21	(0.06, 0.59)	0.12
4-5	0.06	(0.03, 0.10)	0.07	(0.04, 0.12)	0.06		0.02	(0.00, 0.17)	0.08

* Assumed underlying baseline probabilities.
SB, systematic biopsy; ISUP grade: International Society of Urological Pathology.

Table 68 Probabilities (median and 95%CrI) of being classified at different ISUP grades for patients with a previous negative biopsy based on the independent effects analysis.

	Artemis probabilities from Filson (2016)[96] previous negative biopsy data					Artemis + SB probabilities from Filson (2016)[96] previous negative biopsy data		
ISUP	**Cognitive**		**SB**		**Artemis***	**Cognitive + SB**		**Artemis + SB***
No cancer	0.74	(0.68, 0.80)	0.63	(0.57, 0.69)	0.69	0.63	(0.38, 0.76)	0.58
1	0.09	(0.07, 0.12)	0.12	(0.09, 0.15)	0.09	0.14	(0.07, 0.21)	0.15
2	0.06	(0.04, 0.08)	0.11	(0.09, 0.14)	0.10	0.05	(0.02, 0.12)	0.12
3	0.07	(0.04, 0.10)	0.08	(0.05, 0.12)	0.08	0.14	(0.04, 0.47)	0.09
4-5	0.04	(0.02, 0.07)	0.05	(0.03, 0.09)	0.05	0.01	(0.00, 0.12)	0.06

* Assumed fixed

Table 69 Model fit statistics for the cancer detection NMAs

	Any cancer				Clinically significant cancer			
	Model 2a		**Model 2b**		**Model 3a**		**Model 3b**	
Model	**ResDev[1]**	**DIC**	**ResDev[2]**	**DIC**	**ResDev[3]**	**DIC**	**ResDev[4]**	**DIC**
NMA random-effects	33.56	55.35	36.60	66.13	23.11	42.76	26.51	49.33
NMA fixed-effect	37.54	54.57	43.21	67.26	40.29	53.32	34.47	52.65
UME random-effects	NA	NA	NA	NA	23.30	43.53	25.83	48.49
UME fixed-effect	37.95	56.96	44.35	71.48	NA	NA	NA	NA

Shaded cells denote preferred model.
[1]Compare to 35 data points, [2]Compare to 37 data points, [3]Compare to 24 data points, [4]Compare to 26 data points.
NA, not applicable; NMA, network meta-analysis, UME, unrelated mean effects, ResDev, residual deviance, DIC, deviance information criteria.

Table 70 Odds ratios (median and 95% CrI) of cancer detection.

		Any cancer (model 2a)				Clinically significant cancer (model 3a)	
Device Y compared to X		**Fixed-effect NMA**		**Random-effects NMA**		**Random-effects NMA**	
X	**Y**	**median**	**(95%CrI)**	**median**	**(95%CrI)**	**median**	**(95%CrI)**
Cognitive	Systematic	1.37	(1.11, 1.68)	1.32	(0.99, 1.70)	1.18	(0.72, 1.89)
Cognitive	Software	1.30	(1.06, 1.61)	1.29	(1.00, 1.67)	1.35	(0.86, 2.10)
Cognitive	Cognitive + SB	1.56	(1.16, 2.12)	1.54	(1.08, 2.16)	2.47	(1.20, 4.98)
Cognitive	Software + SB	2.05	(1.60, 2.61)	2.03	(1.49, 2.75)	2.71	(1.56, 4.71)
Systematic	Software	0.95	(0.82, 1.11)	0.98	(0.81, 1.24)	1.15	(0.80, 1.66)
Systematic	Cognitive + SB	1.15	(0.86, 1.53)	1.17	(0.84, 1.66)	2.09	(1.07, 4.10)
Systematic	Software + SB	1.50	(1.27, 1.77)	1.54	(1.24, 1.99)	2.29	(1.56, 3.52)
Software	Cognitive + SB	1.20	(0.88, 1.63)	1.19	(0.82, 1.70)	1.82	(0.90, 3.64)
Software	Software + SB	1.57	(1.32, 1.86)	1.57	(1.25, 1.98)	2.00	(1.34, 3.07)
Cognitive + SB	Software + SB	1.31	(0.96, 1.78)	1.32	(0.92, 1.91)	1.10	(0.56, 2.22)

CrI, credible interval; NMA, network meta-analysis; SB: systematic biopsy.

Table 71 All pairwise odds ratios (median and 95% CrI) of any cancer detection (model 2b),

Device Y compared to X		**Fixed-effect NMA**		**Random-effects NMA**	
X	**Y**	**median**	**(95% CrI)**	**median**	**(95% CrI)**
Cognitive	SB	1.39	(1.11, 1.73)	1.31	(0.92, 1.78)
Cognitive	Artemis	1.24	(0.98, 1.58)	1.20	(0.81, 1.75)
Cognitive	Biojet	2.49	(1.47, 4.27)	2.43	(1.08, 5.24)
Cognitive	BiopSee	1.26	(0.67, 2.38)	1.26	(0.56, 2.81)
Cognitive	Urostation	1.45	(1.05, 2.01)	1.41	(0.88, 2.22)
Cognitive	Uronav	1.55	(0.93, 2.62)	1.47	(0.67, 3.06)
Cognitive	Cognitive + SB	1.56	(1.15, 2.13)	1.53	(1.01, 2.30)
Cognitive	Artemis + SB	2.00	(1.51, 2.65)	2.01	(1.23, 3.33)
Cognitive	Urostation + SB	2.18	(1.51, 3.13)	2.10	(1.23, 3.49)
Cognitive	Uronav + SB	2.35	(1.37, 4.07)	2.24	(1.00, 4.77)
SB	Artemis	0.90	(0.74, 1.09)	0.92	(0.64, 1.36)
SB	Biojet	1.80	(1.08, 3.00)	1.85	(0.85, 4.06)
SB	BiopSee	0.91	(0.47, 1.78)	0.96	(0.41, 2.35)
SB	Urostation	1.04	(0.79, 1.38)	1.08	(0.73, 1.63)
SB	Uronav	1.12	(0.70, 1.79)	1.12	(0.57, 2.23)
SB	Cognitive + SB	1.13	(0.84, 1.51)	1.17	(0.80, 1.77)
SB	Artemis + SB	1.44	(1.16, 1.79)	1.53	(1.01, 2.52)

SB	Urostation + SB	1.57	(1.15, 2.13)	1.60	(1.04, 2.50)
SB	Uronav + SB	1.69	(1.04, 2.80)	1.71	(0.85, 3.46)
Artemis	Biojet	2.01	(1.25, 3.22)	2.01	(1.01, 3.98)
Artemis	BiopSee	1.02	(0.52, 2.00)	1.05	(0.43, 2.58)
Artemis	Urostation	1.17	(0.84, 1.62)	1.17	(0.69, 1.97)
Artemis	Uronav	1.25	(0.76, 2.08)	1.22	(0.55, 2.63)
Artemis	Cognitive + SB	1.26	(0.91, 1.74)	1.27	(0.79, 2.07)
Artemis	Artemis + SB	1.61	(1.29, 2.01)	1.66	(1.06, 2.76)
Artemis	Urostation + SB	1.75	(1.22, 2.50)	1.75	(0.99, 3.04)
Artemis	Uronav + SB	1.89	(1.12, 3.24)	1.86	(0.83, 4.07)
Biojet	BiopSee	0.51	(0.22, 1.16)	0.52	(0.17, 1.62)
Biojet	Urostation	0.58	(0.32, 1.03)	0.58	(0.25, 1.39)
Biojet	Uronav	0.62	(0.31, 1.25)	0.61	(0.22, 1.71)
Biojet	Cognitive + SB	0.63	(0.35, 1.11)	0.63	(0.28, 1.48)
Biojet	Artemis + SB	0.80	(0.47, 1.35)	0.83	(0.37, 1.97)
Biojet	Urostation + SB	0.87	(0.48, 1.58)	0.87	(0.36, 2.12)
Biojet	Uronav + SB	0.94	(0.46, 1.93)	0.93	(0.32, 2.62)
BiopSee	Urostation	1.15	(0.56, 2.32)	1.11	(0.44, 2.81)
BiopSee	Uronav	1.23	(0.54, 2.80)	1.16	(0.38, 3.41)
BiopSee	Cognitive + SB	1.24	(0.61, 2.49)	1.21	(0.49, 3.01)
BiopSee	Artemis + SB	1.58	(0.79, 3.13)	1.59	(0.63, 4.14)
BiopSee	Urostation + SB	1.73	(0.83, 3.56)	1.66	(0.63, 4.26)
BiopSee	Uronav + SB	1.86	(0.80, 4.30)	1.78	(0.57, 5.26)
Urostation	Uronav	1.07	(0.62, 1.86)	1.04	(0.47, 2.28)
Urostation	Cognitive + SB	1.08	(0.73, 1.60)	1.09	(0.64, 1.86)
Urostation	Artemis + SB	1.38	(0.97, 1.95)	1.42	(0.81, 2.62)
Urostation	Urostation + SB	1.50	(1.10, 2.05)	1.49	(0.96, 2.29)
Urostation	Uronav + SB	1.62	(0.92, 2.88)	1.59	(0.70, 3.51)
Uronav	Cognitive + SB	1.01	(0.58, 1.74)	1.04	(0.48, 2.34)
Uronav	Artemis + SB	1.29	(0.76, 2.14)	1.36	(0.62, 3.20)
Uronav	Urostation + SB	1.40	(0.80, 2.45)	1.42	(0.64, 3.25)
Uronav	Uronav + SB	1.51	(0.92, 2.51)	1.52	(0.75, 3.07)
Cognitive + SB	Artemis + SB	1.28	(0.92, 1.76)	1.31	(0.81, 2.20)
Cognitive + SB	Urostation + SB	1.39	(0.92, 2.11)	1.37	(0.76, 2.42)
Cognitive + SB	Uronav + SB	1.50	(0.85, 2.69)	1.46	(0.64, 3.23)
Artemis + SB	Urostation + SB	1.09	(0.75, 1.58)	1.05	(0.55, 1.88)
Artemis + SB	Uronav + SB	1.18	(0.69, 2.03)	1.12	(0.47, 2.50)
Urostation + SB	Uronav + SB	1.08	(0.61, 1.94)	1.07	(0.46, 2.43)

Table 72 All pairwise odds ratios (median and 95% CrI) of clinically significant cancer detection (model 3b)

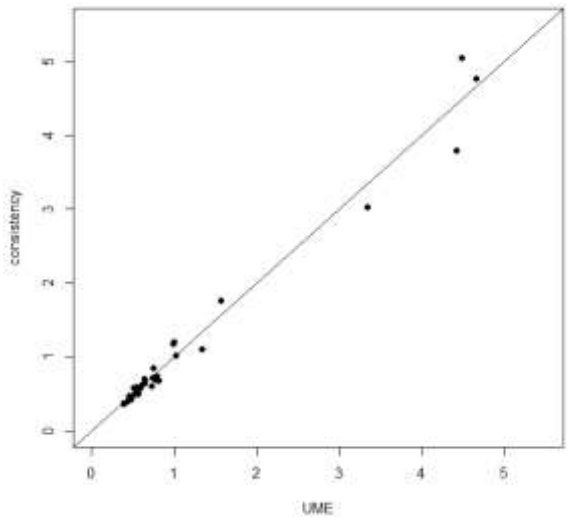
Device Y compared to X		Random-effects NMA	
X	**Y**	**median**	**(95% CrI)**
Cognitive	SB	1.30	(0.71, 2.24)
Cognitive	Artemis	1.44	(0.80, 2.47)
Cognitive	Biojet	4.79	(1.56, 14.56)
Cognitive	BiopSee	1.04	(0.38, 2.85)
Cognitive	Urostation	1.65	(0.72, 3.64)
Cognitive	Cognitive + SB	2.41	(1.10, 5.18)
Cognitive	Artemis + SB	2.32	(1.13, 5.28)
Cognitive	Urostation + SB	3.71	(1.55, 7.91)
SB	Artemis	1.10	(0.65, 1.90)
SB	Biojet	3.69	(1.23, 11.40)
SB	BiopSee	0.80	(0.26, 2.61)
SB	Urostation	1.27	(0.72, 2.27)
SB	Cognitive + SB	1.86	(0.89, 3.91)
SB	Artemis + SB	1.78	(0.98, 3.78)
SB	Urostation + SB	2.85	(1.56, 4.94)
Artemis	Biojet	3.34	(1.28, 8.88)
Artemis	BiopSee	0.72	(0.23, 2.34)
Artemis	Urostation	1.15	(0.52, 2.53)
Artemis	Cognitive + SB	1.68	(0.75, 3.75)
Artemis	Artemis + SB	1.62	(0.86, 3.50)
Artemis	Urostation + SB	2.59	(1.12, 5.45)
Biojet	BiopSee	0.22	(0.05, 0.99)
Biojet	Urostation	0.34	(0.10, 1.18)
Biojet	Cognitive + SB	0.51	(0.14, 1.77)
Biojet	Artemis + SB	0.48	(0.15, 1.70)
Biojet	Urostation + SB	0.77	(0.21, 2.59)
BiopSee	Urostation	1.59	(0.43, 5.71)
BiopSee	Cognitive + SB	2.32	(0.64, 8.11)
BiopSee	Artemis + SB	2.24	(0.66, 8.19)
BiopSee	Urostation + SB	3.56	(0.92, 12.33)
Urostation	Cognitive + SB	1.47	(0.57, 3.75)
Urostation	Artemis + SB	1.41	(0.62, 3.66)

Urostation	Urostation + SB	2.25	(1.23, 3.86)
Cognitive + SB	Artemis + SB	0.96	(0.47, 2.23)
Cognitive + SB	Urostation + SB	1.53	(0.58, 3.80)
Artemis + SB	Urostation + SB	1.60	(0.59, 3.53)

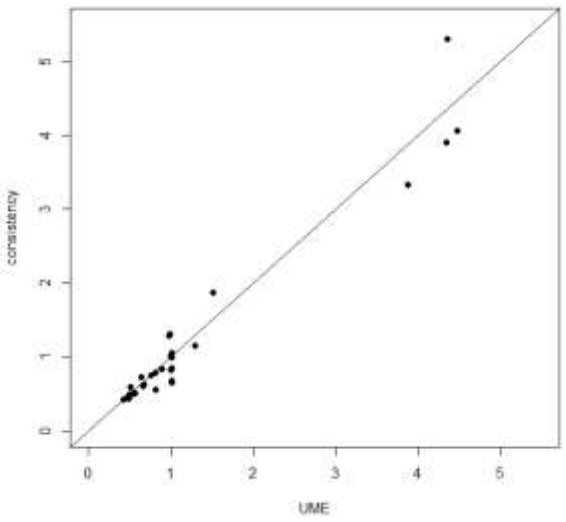
Results from additional analyses: Figures

Figure 14 Plots of residual deviance contributions for the NMA (consistency) and unrelated mean effects model.

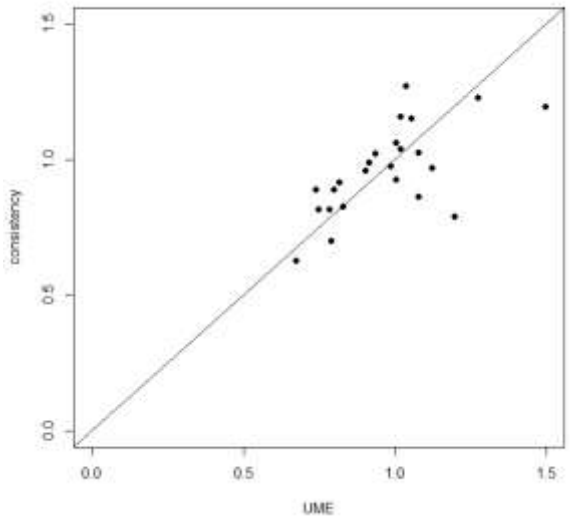
a) Any cancer, fixed-effect NMA (model 2a)



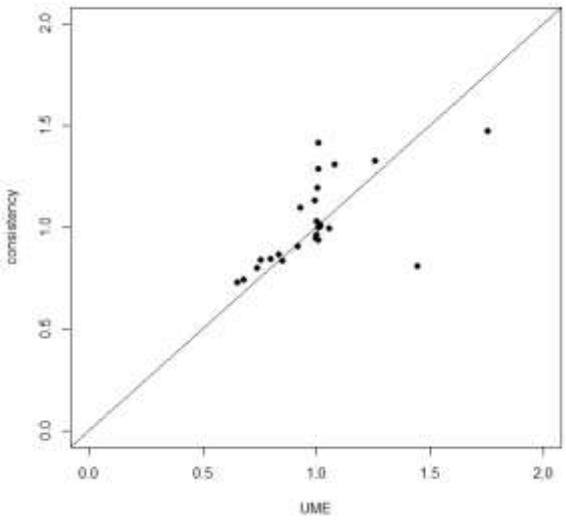
b) Any cancer, fixed-effect NMA (model 2b)



c) Clinically significant cancer, random-effects NMA (model 3a)



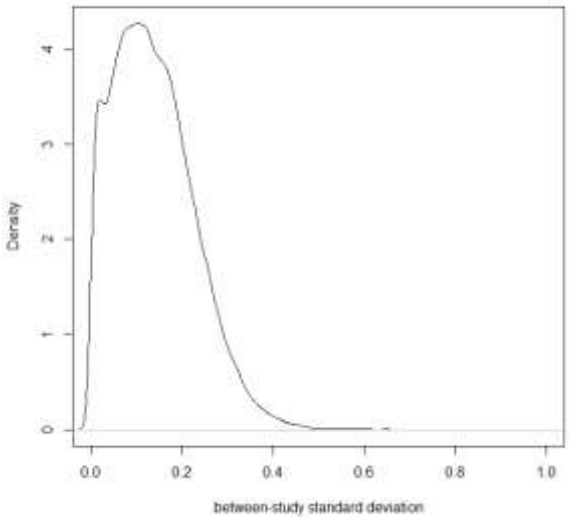
d) Clinically significant cancer, random-effects NMA (model 3b)



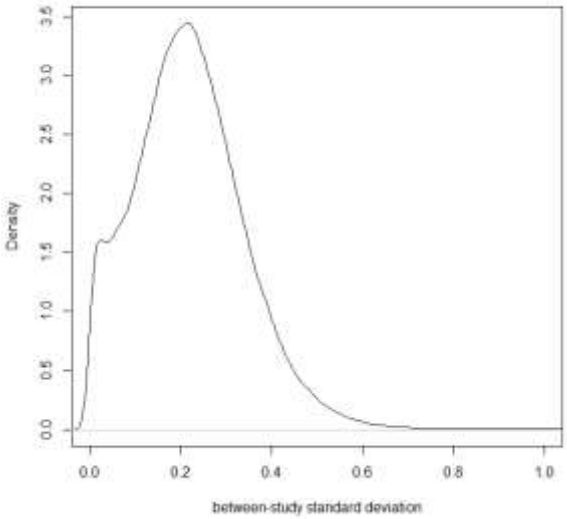
UME, unrelated mean effects

Figure 15 Posterior densities of the between-study standard deviation for random-effects models.

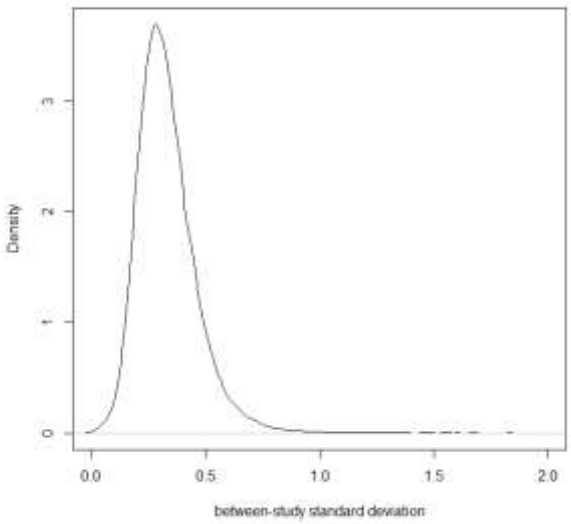
a) Any cancer (model 2a)



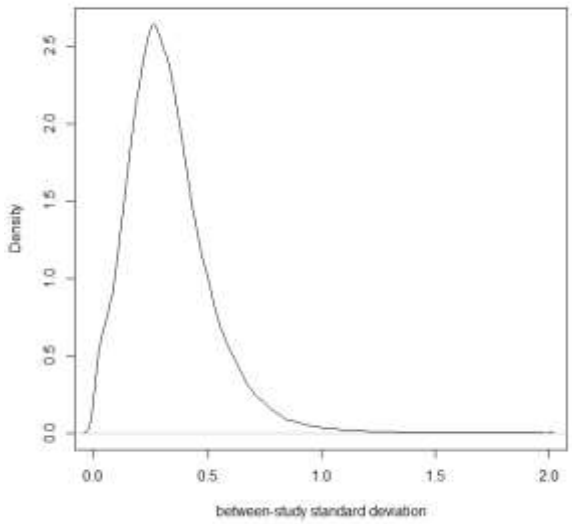
b) Any cancer (model 2b)



c) Clinically significant cancer (model 3a)



d) Clinically significant cancer (model 3b)



APPENDIX 7. ADDITIONAL RESULTS FROM STUDIES INCLUDED IN THE SYSTEMATIC REVIEW

Table 73 Additional diagnostic accuracy results from studies not included in the meta-analyses

Study	Design	Pop.	Tests	N	Outcome (Metric)	Summary	Effect estimates	Direction of effect/p-value
Delongchamps (2013)[98]	Consecutive series, between-patient	BN	SF (Koelis - Urostation Touch) vs SB*	82	CsPCa: Gleason ≥3+4 (N)	SF alone detected 35 of the 44 cancers detected by SB as well as 27 undetected by SB, of which 8 had a Gleason score of greater than 6. All 9 cancers missed by SF but detected by SB had a Gleason score of 6, of which 7 involved less than 5 mm of the biopsy core.	SF: 27; CF: 18	Favours SF vs SB p=0.01
			CB vs SB	54	PCa (OR) Definition NR	CF alone detected 37 of the 55 cancers detected by SB as well as 3 undetected by RB, of which 2 had a Gleason score of greater than 6. Of the 18 cancers missed by SF but detected by RB, 16 had a Gleason score of 6 and 15 involved less than 5 mm of the biopsy core. Conditional logistic regression analysis showed that cognitive fusion was not significantly better at detecting prostate cancer compared to systematic biopsy (OR not reported)	NR	No significant difference (p = 0.66)
Hansen (2018)[95]	Prospective, between patients	BN	SF vs CF; SB vs SF+CF+SB	SF: 395 CF: 176	Pca (PI-RADS 3)	Favours combination biopsy over targeted biopsy alone. No significant difference between combination biopsy and systematic biopsy	SB: 53% CF+SF: 38% SB+SF+CF: 56%	p <0.001 (TB vs SB+TB) P = 0.063 (SB vs SB+TB)
					Pca (PI-RADS 4-5)	Favours combination biopsy over systematic biopsy or targeted biopsy alone	SB: 80% CF+SF: 73% SB+SF+CF: 88%	p < 0.001 (both)

					CsPCa: Gleason ≥3+4 (PI-RADS 3)	Favours combination biopsy over targeted biopsy alone. No significant difference between combination biopsy and systematic biopsy	SB: 37% SF+CF: 21% SB+SF+CF: 30%	p <0.001 (TB vs SB+TB) p = 0.125 (SB vs SB+TB)
					CsPCa: Gleason ≥3+4 (PI-RADS 4-5)	Favours combination biopsy over systematic biopsy or targeted biopsy alone	SB: 61% CF+SF: 59% SB+SF+CF: 71%	p < 0.001 (both)
Ferriero (2022)[81]	Prospective cohort, between patients	BN + RB	SF (Urostation) vs SF (Biojet)	Urostation: 103 Biojet: 211	Pca per target (%) Definition NR	No significant differences between the two software fusion types	SF (Urostation): 69.8%, SF (Biojet): 56.6%	Not significant p = 0.077
					CsPCa per (%) Definition NR	No significant differences between the two software fusion types	SF (Urostation): 50.6%, SF (Biojet): 50.6%	Not significant p = 1.0
Sokolakis (2021)[83]	Prospective cohort, between patients	BN + RB	SF (Biojet) vs SF (Trinity) vs SF (Uronav)	Biojet: 20 Trinity: 20 Uronav: 20	ISUP 1 (N)	No significant difference between the three software types	Biojet: 2, Trinity: 3, UroNav: 3	No significant difference. P > 0.99
					ISUP 2 (N)		Biojet: 4, Trinity: 4, UroNav: 4	
					ISUP 3 (N)		Biojet: 4, Trinity: 3, UroNav: 3	
					ISUP 4 (N)		Biojet: 1, Trinity: 2, UroNav: 2	
					ISUP 5 (N)		Biojet: 1 , Trinity: 2, UroNav: 1	
Liang (2020)[85]	Retrospective cohort, between patients	BN	SF (Predictive Fusion Software) vs CF	SF: 92 CF: 71	ISUP 1 (%)	Similar detection rates (within 5%)	SF = 17%, CF = 21%	Significance NR
					ISUP 2 (%)		SF = 14%, CF = 13%	
					ISUP 3 (%)		SF = 9%, CF = 11%	
					ISUP 4 (%)		SF = 8%, CF= 13%	
					ISUP 5 (%)		SF = 3%, CF = 3%	
Lockhart (2022)[100]	Retrospective cohort, between patients	BN, AS	SF (MIM Fusion Software), vs CF	SF: 131 CF: 223	ISUP 2	Multinomial logistic regression analysis was performed to explore potential factors affecting csPCa detection rates. Fusion or	NR	p=0.729

						cognitive biopsy made no difference to CsPCa detection rates		
Monda (2018)[90]	Retrospective cohort, between patients	BN + RB	SF (UroNav) vs CF vs SB (concurrent)	SF/SB: 162 CF/SB: 348	Gleason 6 (%)	Higher rate of Pca detection with cognitive targeted biopsy	SF: 14.4%, CF: 22.8%	Significance NR
					Gleason 7 (%)	Similar rates of detection	SF: 20.1%, CF: 18.5%	Significance NR
					Gleason 8 (%)	Similar rates of detection	SF: 3.4%, CF: 3.1%	Significance NR
					Gleason 9-10 (%)	Similar rates of detection	SF: 4.3%, CF: 5.6%	Significance NR
					Missed targeted biopsy (%) TB <7 & SB >7	Similar rates (within 5%)	SF: 5.5%, CF: 9.9%	Not significant p=0.172
					Equivalent (%) TB & SB ≥7 or TB & SB < 7		SF: 85.1%, CF: 82.1%	Not significant p=0.172
					Upstage (%) TB ≥7 & SB < 7		SF: 9.5%, CF: 8.0%	Not significant p=0.172

* Also compared to Esaote rigid software fusion system.

BN, Biopsy naïve; RB, repeat biopsy; AS, active surveillance; SF, software fusion; CF, cognitive fusion; SB, systematic biopsy; CsPCa, clinically significant prostate cancer; PCa, prostate cancer; PI-RADS, Prostate Imaging Reporting & Data System; ISUP, International Society of Urological Pathology; NR, not reported; OR, odds ratio

Table 74 Test positive rates for patients undergoing repeat biopsy following prior negative biopsy

Study	Population	SF technology	Route#	Anaesthesia#	N of patients	Outcome (definition)	Biopsy positive rates	Statistical significance
FUTURE (2019)[31]	Prior negative SB within median 8 months (IQR 4–23)	Biopsee	SF: TP CF: TR	SF: GA CF: LA	SF: 79 CF: 78	PCa (NR)	SF: 49.4% CF: 43.6%	p=0.4
						CsPCa (GS: ≥3+4)	SF: 34.2% CF: 33.3%	p>0.9
PROFUS (2014)[97]	Prior negative biopsy (no further details)	Artemis	TR	LA	SF & CF*: 34	PCa (NR)	SF: 29.4% CF: 23.5%	NR
						CsPCa (GS: ≥3+4)	SF: 20.6% CF: 14.7%	NR

For SF and CF approaches unless otherwise specified; * within-patient comparison; PCa: prostate cancer; CsPCa: clinically significant prostate cancer; SF: Software fusion; CF: cognitive fusion; SB, systematic biopsy; NR: not reported; GS: Gleason score; TR, transrectal; TP, transperineal; LA, local anaesthetic; GA, general anaesthetic.

Table 75 Test positive rates for biopsy naïve patients

Study	SF technology	Route#	Anaesthesia#	N of patients	Outcome (definition)	Test positive rate	Statistical significance
Delongchamps (2013)[98]	Urostation Touch (Koelis)*	TR	NR	SF: 82 CF: 54	**PCa** (NR)	NR&	SF vs. SB: p=0.006 CF vs. SB: p=0.22
					csPCa (NR)	NR&	SF vs. SB: p=0.001 CF vs. SB: p=0.6
Ferriero (2022)[81]	Urostation; Biojet	Urostation: TR; Biojet: NR	NR	Urostation: 103 Biojet: 232 (1:1 PS matched cohort, n = 83)	**PCa (**GS 6)	Urostation: 69.8% Biojet: 56.6%	p = 0.077
					csPCa (GS ≥ 7)	Urostation: 50.6% Biojet: 50.6%	p = 1
Hansen (2018)[95]	Biopsee	TP	GA	SF: 395 CF: 176	**PCa** (NR)	SF: 53% CF: 38%	NR
					csPCa ()	SF: 56% CF: 70%	NR
Izadpanahi (2021)[82]	Artemis	TR	LA	SF: 99 CF: 100	**PCa** (GS 6 and <4mm core length)	SF: 44.4% CF: 31.0%	p = 0.035
					csPCa (GS ≥ 7 or GS 6 and ≥4mm core length)	SF: 33.3% CF: 19.0%	p = 0.016
Liang (2020)[31, 85]	BK Fusion^	TP	LA	SF: 92 CF: 71	**PCa** (GS 6)	SF: 51.08% CF: 60.56%	p=0.228

					csPCa ()	SF: 35.87% CF: 39.43%	p=0.641
Lockhart (2022)[100]	BK fusion^	TP	NR	SF+SB: 97 CF+SB: 186	**csPCa** (GS ≥ 7)	SF+SB: 53% CF+SB: 66.7%	NR
PAIREDCAP (2019)[88]	Artemis	TR	LA	248	**PCa** (GS 6)	SF: 17.3% CF: 15.3%	NR
					csPCa (GS ≥ 7)	SF: 54.0% CF: 46.8%	NR
PROFUS (2014)[97]	Artemis	TR	LA	67	**PCa** (GS 6)	SF: 35.8% CF: 34.3%	NR
					csPCa (GS ≥ 7)	SF: 28.4% CF: 26.9%	NR

[&] Probability of detecting cancer undetected by SB against SB as reference was calculated but NR # For SF and CF approaches unless otherwise specified; ^ 'Predictive Fusion Software', * Also compared to Esaote rigid software fusion system

PCa: prostate cancer; CsPCa: clinically significant prostate cancer; SF: Software fusion; CF: cognitive fusion; SB, systematic biopsy; NR: not reported; GS: Gleason score; TR, transrectal; TP, transperineal; LA, local anaesthetic; GA, general anaesthetic.

APPENDIX 8. STUDIES OF LONG-TERM MORBIDITY AND MORTALITY OUTCOMES

Table 76 Long-term outcome studies: radical radiotherapy

Study	Design	Population	Treatment	Comparator	Outcome
ACENDE-RT [113, 177, 178]	RCT N=398	Intermediate-high risk. CPG 4-5	Low dose rate brachytherapy+external beam radiotherapy	dose-escalated external beam radiation therapy	Local recurrence, distant metastases, OS (KM). F-u up to 10yrs
HYPRO [179 114]	RCT, n=820	Intermediate-high risk	hypofractionated radiotherapy	Conventional radiotherapy	OS, 7-yr relapse free survival, AE F-u up to 10yrs
PROFIT [115]	RCT, n=1206	Intermediate	hypofractionated radiotherapy	Conventional radiotherapy	OS, biochemical failure, AE, f-u up to 5yrs HRQOL-48 weeks
CCHiP [116, 180]	RCT, n=3216	Intermediate-high risk	hypofractionated radiotherapy	Conventional radiotherapy	OS, relapse free survival, AE F-u up to 8 yrs
HYPO-RT-PC [117, 181]	RCT N=1180	Intermediate-high risk	ultra-hypofractionation	conventional fractionated radiotherapy	Failure free survival and PCa-specific survival (5yr) QoL (6yrs)
Marzi 2009[118]	RCT, n=162	Intermediate-high risk Gleason 7-10	hypofractionated radiotherapy	Conventional radiotherapy	OS. f-u 30 months

Table 77 Long-term outcome studies: Radiotherapy+androgen deprivation therapy vs. radiotherapy alone

Study	Design	Population	Treatment	Comparator	Outcomes
Kishan, (2022)[121]	IPD M-A	Intermediate-high risk	Radiotherapy+androgen deprivation therapy (incl. as prolongation therapy)	Radiotherapy alone	Metastasis-free survival (KM) OS (KM). 11.4 yrs f-u. Biochemical recurrence, distant metastasis.

Table 78 Long-term outcomes studies: Prostatectomy vs. observation

Study	Design	Population	Treatment	Comparator	Outcomes
PIVOT[182] [119]	RCT	Low, intermediate and high	Radical prostatectomy	Watchful waiting	OS, PCa death, distant metastases, AEs f-u 22.1yrs
SPCG4[120] [183]	RCT	Localised, non-metastatic	Radical prostatectomy	Watchful waiting	Overall mortality, PCa death, distant metastases, AEs, QoL F-u: 29 yrs

Table 79 Long-term outcomes studies: Radical prostatectomy vs radical radiotherapy vs. observation

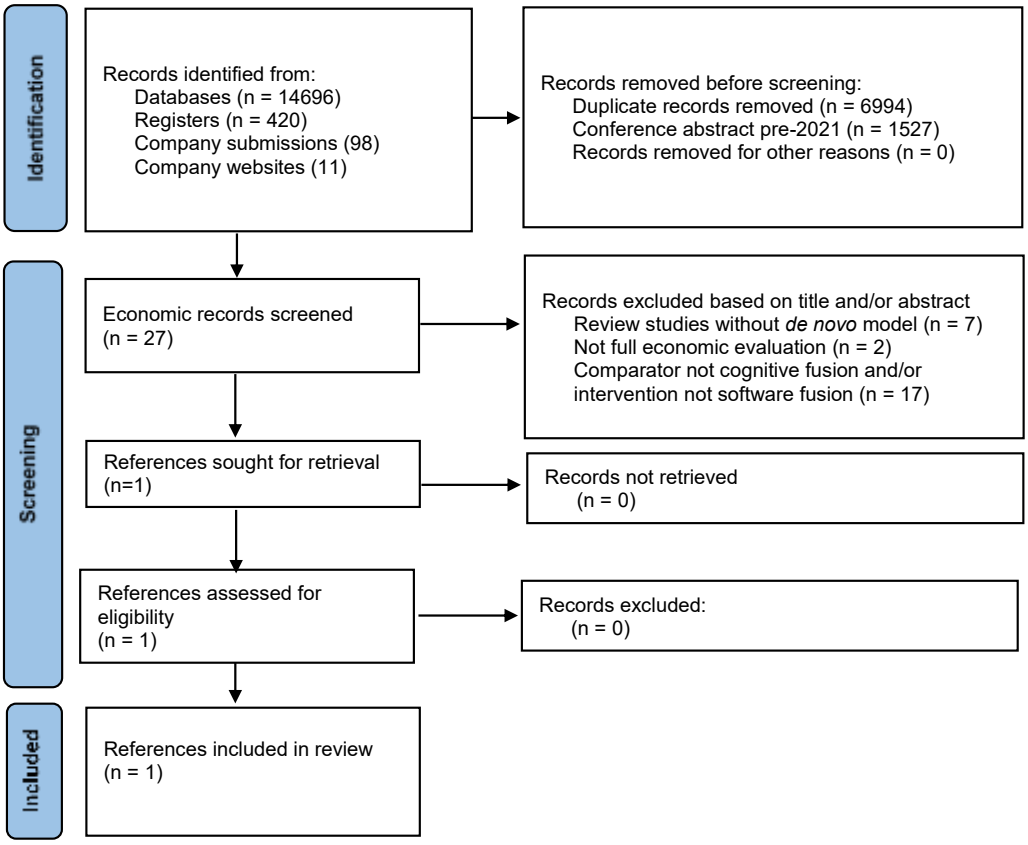
Study	Design	Population	Treatment	Comparator	Outcomes
PROTecT [124 55]	RCT N=1643	Localised, non-metastatic	Radical prostatectomy, , radical radiotherapy	AS	PFS, patient-centred outcomes F-u: median 10yrs

Table 80 Long-term outcomes studies: Docetaxel & hormone sensitive therapy

Study	Design	Population	Treatment	Comparator	Outcomes
STAMPEDE[59] [151]	RCT, n=1776	High risk PCa (Gleason 8-10) and metastatic	Androgen deprivation therapy plus docetaxel and estramustine	androgen deprivation therapy alone	OS, PFS. F-u: 6.5yrs
GETUG 12[60] [184]	RCT, n=413	High risk PCa (Gleason 8-10)	Addition of docetaxel, zoledronic acid/estramustine, or both to first-line long-term hormone therapy	Long-term hormone therapy	OS, PFS F-u: 12 yrs
TAX-3501[61]	RCT N=228	Metastatic, post-radical prostatectomy	Docetaxel and leuprolide	Leuprolide alone	OS, PFS, AEs, f-u 3.4 yrs

APPENDIX 9. REVIEW OF COST-EFFECTIVENESS EVIDENCE

Figure 16 PRISMA flow diagram for cost-effectiveness of software fusion systems review



Critical appraisal of cost-effectiveness studies of MRI Fusion systems Pahwa et al., 2017

Table 81 Yang et al., 2019[125] Checklist for model-based economic evaluations of diagnostic tests: Pahwa et al., 2017[127]

	Response (Y,N or NA)	Comments
1. Decision problem and scope specified		
Is there a clear statement of the decision problem?	Y	
Is the perspective of the model stated clearly?	N	
Has the target population been identified?	Y	
Are the model inputs consistent with the stated perspective?	NA	Perspective not stated clearly
Are the primary outcomes of the model consistent with the perspective, scope and overall objective of the model?	NA	Perspective not stated clearly
2. Identification and description of the comparators		
Have all the feasible and practical options been identified?	Unclear	Authors do not state whether there are other feasible and relevant alternatives
Have the comparators being evaluated been clearly described?	Y	
If comparators have been excluded from the evaluation, have these exclusions been justified?	NA	
3. Appropriate data identification		
Are the data identification methods transparent, systematic and appropriate given the objectives of the model?	N	The data identification methods are not described
4. Sufficient detail for data incorporation		
Have all data incorporated into the model been described and referenced in sufficient detail?	Y	
Where choices have been made between data sources, are these justified appropriately?	N	
Are transition probabilities calculated appropriately?	NA	Not a state transition model
Has discounting been conducted?	Y	

5. Quality and incorporation of test accuracy data		
Has the quality of the test accuracy data been assessed?	N	
Have diagnostic accuracy data been derived from high quality data sources (hierarchy of evidence)?	NA	Sources of data to inform data accuracy are not described in sufficient detail to establish quality of data
Are tests in sequence treated dependently, where appropriate?	N	Dependencies between tests in a sequence not modelled (implicit assumption of independence between tests in each sequence)
6. Quality and incorporation of treatment data		
Has the quality of the treatment effect data been assessed?	N	Linkage to long-term outcomes is done via lifetime pay-offs applied to diagnostic decision tree – relative treatment effects are not applied in the model
Have relative treatment effects been derived from high quality data sources (hierarchy of evidence)?	NA	
7. Source and incorporation of cost data		
Has the source of cost data been presented clearly?	Y	
Have costs been inflated to a specific year, where appropriate?	Y	
8. Source and incorporation of utility data		
Is the source for the utility weights referenced and justified?	N	Assumption that 1 LY corresponds to 1 QALY in healthy individuals (no prostate cancer) is not supported by empirical data.
Are the utilities incorporated into the model appropriately?	Unclear	Most QALYs are estimated directly from an external Markov model
9. Model structure		
Have the reasons behind the type of decision analytic model chosen been fully described and justified?	N	
Has a systematic review of existing economic evaluations been carried out?	N	
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	NA	The structure of the model is not sufficiently described or depicted to assess whether it is consistent with the health condition.
Are the structural assumptions underpinning the model transparent and justified?	Partly	Not all assumptions are justified, and some assumptions are not explicit (e.g., independence between results of tests in a sequence)
Have the methods used to extrapolate short-term results to final outcomes been documented and justified?	NA	Linkage to long-term outcomes is done via lifetime pay-offs applied to diagnostic decision tree
Has the time horizon been stated and justified?	Y	
Has cycle length of Markov models been justified?	NA	Not a Markov model
10. Uncertainty		

Has parameter uncertainty been addressed via sensitivity analysis?	Y	One-way sensitivity analysis
Has probabilistic sensitivity analysis been carried out? If not, has this omission been justified?	Y	
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	Partly	The ranges used are not clearly justified for most parameters
If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	N	Probability distributions for each parameter are not described
Have structural uncertainties been addressed via sensitivity analysis?	N	
Have alternative assumptions related to final outcomes been explored through sensitivity analysis?	N	
Has value of information analysis been done?	N	
11. Validity		
Has the face validity been reviewed by someone external to the model developers?	N	Not described
Has the mathematical logic of the model been assessed? (e.g., using null and extreme values)	N	Not described
Have the model and its results been compared to the findings of other models and studies, and any disagreements or inconsistencies been explained (cross-validity)?	Y	

LY, life-year; N, no; NA, not applicable; Y, yes.

Results of the additional targeted reviews to support model conceptualisation

The searches described in Section 4.1 identified twenty-seven titles of which sixteen did not meet the inclusion criteria based on title and/or abstract. Full text publications were obtained for the remaining ten records.[129, 130, 132, 134-136, 139-142] In addition, economic studies were identified from a systematic review in a previous DAR by the Southampton EAG[126]. We identified five additional publications from the previous DAR economic evidence review.[131, 133, 137, 138, 143]

In total, sixteen titles comprising fifteen cost-effectiveness models[126, 129-143] were considered potentially relevant to inform the *de novo* model conceptualisation for inclusion. We note that the Wilson et al. (2021)[131] model is structurally similar (and shares many common evidence sources) to the cost-effectiveness model developed in the context of the PROMIS trial[135, 136] (henceforth referred to as the PROMIS model), although it does not model the full range of strategies in PROMIS. Similarly, the Southampton DAR model[126] is an extension of the model developed in the context of the 2019 update of the NICE CG131[133] (henceforth referred to as the NICE CG131 model). These studies are summarised in Table 82.

Table 82 Studies identified as potentially relevant to inform model conceptualisation

Study:1st author (year), country **Study aim**	**Diagnostic strategies**	**Definition of CS PCa**	**Biopsy diagnostic accuracy outcomes**	**Model structure and modelling approach**	**Evidence linkage to long-term outcomes**
Souto-Ribeiro (2022)[126], UK To assess the CE of LATP vs. LATRUS and GATP in men with suspected PCa for whom prostate biopsy is indicated	1st biopsy (+): treated, (-): % discharged/monitored & % 2nd biopsy; →2nd biopsy (-): discharged/monitored; 2nd biopsy (+): treated Biopsy options: 1st Biopsy - LATP (w/wo specific freehand devices)/ GATP/ LATRUS; 2nd biopsy - LATRUS	Histopathological definition: Gleason score ≥ 4 and/or a cancer core length≥ 4 mm Clinical definition: - CNS PCa: Gleason score≤ 6, PSA ≤10 ng/ml & T1-T2a stage (=LR) - CS PCa: Gleason=7, PSA 10-20 ng/ml & T2b stage (=IR); or Gleason score≥8, PSA>20 ng/ml & ≥T2c stage (=HR)	Probabilities of TRUS detecting CNS & CS PCa (stratified by LR, IR, HR) RRs for PCa detection rates for LATP & GATP vs. LATRUS are applied to baseline probabilities (with LATRUS) Specificity of detecting PCa	Decision Tree: classifies patients according to diagnostic accuracy, true disease status & underlying risk category. Tree also captures biopsy complications + Markov model capturing treatment allocation conditional on classification & longer-term outcomes Health states: No PCa; unDx LR; unDx IR; unDx HR; unDx metastatic; Dx LR; Dx IR; Dx HR; Dx metastatic; PCa death; other cause death.	Via Markov model capturing sequential disease progression from lower to higher risk category (LR→IR→HR) of localised disease and from HR to metastatic disease. PCa mortality only applies to metastatic disease.
Wilson (2021)[131], UK To assess the CE of LATP vs LATRUS for men at risk of PCa who are referred to 2ary care investigations	mpMRI (No/CNS PCa): discharged/monitored; mpMRI (CS PCa); →1st biopsy (CS PCa): treated; 1st biopsy (No/CNS PCa) →2nd biopsy (No/CNS PCa): discharged/monitored; 2nd biopsy (CS PCa): treated Biopsy alternatives: LATP or LATRUS	Histopathological definition: NR Clinical definition: Text suggests LR is equivalent to CNS PCa, & IR/HR to CS PCa, but the risk categories are not defined.	Probabilities of detecting No PCa, LR, IR or HR conditional on true disease status & previous test results (mpMRI/ biopsy) Specificity of detecting PCa	Decision Tree: classifies patients according to diagnostic accuracy, true disease status & underlying risk category. Tree also captures biopsy complications and treatment allocation. + Markov model capturing longer-term outcomes Health states: no PCa (?); progression-free, metastatic disease, death	Via Markov model capturing disease progression from localised disease to metastatic disease. PCa mortality only applies to metastatic disease.

Cheng (2021)[129], Singapore To assess the CE of diagnostic strategies involving combined biopsy in sequences (with/wo SBx/ TPMB) for men with suspected PCa based on elevated PSA and/or abnormal DRE	Test sequence in each strategy: 1. Combined biopsy 2. Combined biopsy→(-): SBx 3. Combined biopsy→(-): TPMB 4. Combined biopsy→(-): SBx→(-): TPMB 5. SBx→(-): Combined biopsy 6. SBx→(-): Combined biopsy→(-): TPMB Where combined biopsy means: . mpMRI → PI-RADS 1,2: % no biopsy & % SBx; PI-RADS 3+: Combined biopsy; .individuals with biopsy(+) receive treatment	Histopathological definition: NR; Clinical definition - CNS PCa: Gleason score <7, PSA <10 ng/ml; & T1-T2a stage (=LR) - CS PCa: Gleason score=7, or PSA 10-20ng/ml; or T2b stage (=IR); or Gleason score>7, PSA >20ng/ml; or ≥ T2c stage (=HR)	For SBx, TBx and combined biopsy: Probabilities of detecting LR, IR, HR conditional on true disease status & prior test results For TPMB: Specificity of detecting PCa, sensitivity to detect LR, IR, HR	Decision Tree: classifies patients according to diagnostic accuracy, true disease status & underlying risk category + Markov model capturing treatment allocation conditional on classification & longer-term outcomes Health states: No PCa, unDx localised PCa, metastatic PCa, correctly Dx localised LR (3 separate treatment health states: WW, AS, RTx ±ADT) localised IR Dx LR (3 separate treatment health states: WW, AS, RTx ±ADT), correctly Dx localised LR (2 separate treatment health states: WW, RTx ±ADT), PCa death, all cause death.	Via Markov model capturing: . 1^{ary} treatment allocation and subsequent treatment changes . Disease progression from localised to metastatic disease PCa mortality only applies to metastatic disease.

Hao (2021), Sweden To assess the CE of diagnostic strategies involving TBx, SBx or combined biopsy for men undergoing (or eligible for) quadrennial PSA screening	1. No PSA screening (assumes average 2 SBx for symptomatic identification) Screening strategies If PSA ≥3ng/mL: 2. SBx 3. mpMRI → PI-RADS<3: rescreening; PI-RADS≥3: TBx 4. mpMRI → PI-RADS <3: rescreening; PI-RADS ≥3: Combined biopsy 5. mpMRI → PI-RADS <3: SBx; PI-RADS ≥3: Combined biopsy Where individuals with biopsy (+) receive treatment, and those with biopsy (-) return to screening	NA	FN rates conditional on the true disease status (ISUP GG1 or GG≥2) Specificity of detecting PCa	Continuous time microsimulation PCa natural history model Health states: .No PCa .Preclinical states: ISUP GG1,T1-T2; ISUP GG1, T3-T4; ISUP GG1 metastatic; ISUP GG2-3,T1-T2; ISUP GG2-3, T3-T4; ISUP GG2-3 metastatic; ISUP GG4-5,T1-T2; ISUP GG4-5, T3-T4; ISUP GG4-5, metastatic .Clinical states: ISUP GG1, ISUP GG1,T1-T2; ISUP GG1, T3-T4; ISUP GG1 metastatic; ISUP GG2-3,T1-T2; ISUP GG2-3, T3-T4; ISUP GG2-3 metastatic; ISUP GG4-5,T1-T2; ISUP GG4-5, T3-T4; ISUP GG4-5, metastatic. .. Diagnosis and treatment submodel for clinical states: diagnosis; localised T1, T2, T3, T4, ISUP GG1 or GG2+ treatment (AS, RP &/or RT, post treatment follow-up), metastatic (treatment, palliative care, terminal illness) .other cause death; PCa death	Via microsimulation model capturing . disease onset and progression from preclinical to clinical PCa. ..Preclinical states reflect disease onset by ISUP GG and progression by T stage to metastatic PCa (from T1-T2→T3-T4→metastatic PCa) ..Disease progression in clinical states seems to be from localised to metastatic .. 1^{ary} treatment allocation and subsequent treatment changes PCa mortality only applies to metastatic disease in clinical states.

Getaneh (2021)[132], The Netherlands To assess the CE of adding mpMRI as triage test between PSA and biopsy for population-based triennial screening	Screening strategies: 1. (not described) PSA screening protocol involving TRUS 2. PSA → PSA <3ng/mL: no further assessment; PSA ≥ 3ng/mL: mpMRI → PI-RADS<3: no biopsy; PI-RADS≥3: TBx Individuals with biopsy (+) receive treatment, and those with biopsy (-) return to screening (not explicit)	NR	TBx: - Sensitivity to detect LG&HG* PCa - Misclassification rate (HG classified as LG) TRUS: - Biopsy sensitivity (not specified whether it applies to PCa or PCa significance) - Misclassification rate (HG classified as LG)	Microsimulation screening analysis; life history model w/wo screening Health states: .No PCa .Preclinical states: T1, GS<7; T1, GS=7; T1, GS>7; T2, GS<7; T2, GS=7; T2, GS>7; T3, GS<7, T3, GS=7; T3, GS>7; each state can be local-regional or distant metastatic (18 health states in total) .Clinical states: T1, GS<7; T1, GS=7; T1, GS>7; T2, GS<7; T2, GS=7; T2, GS>7; T3, GS<7, T3, GS=7; T3, GS>7; each state can be local-regional or distant metastatic; death (18 health states in total plus death)	Via microsimulation model capturing .disease onset and progression from preclinical to clinical PCa by screening or clinical diagnosis ..preclinical states reflects disease onset at T1-GS<7 or T1-GS>7; then progression by T stage (T1→T2→T3→T4) and GS (GS<7→GS=7→GS>7); any state can progress from local-regional state to distant state ..Disease progression in clinical state is not modelled .. 1ary treatment allocation based on age, T stage, GS PCa mortality only applies at clinical states.
NICE (2019)[133], UK To assess the CE of follow-up protocols for people who have a raised PSA, MRI(-) and/ or (-) biopsy	Alternative follow-up protocols, defined according to: .Type of screening test and the related threshold (e.g., PSA derivatives); .Frequency of the screening test; .Type of biopsy if the previous test positive (e.g., TRUS or TPMB); .Stopping rule - defines the duration of follow-up for each strategy.	Histopathological definition: Gleason score ≥3+4 or cancer core length ≥4mm Clinical definition: - CNS PCa: Gleason scor <7 or cancer core length<4mm or PSA≤ 10ng/mL (=LR) - CS PCa: Gleason score=7 or cancer core length≥4mm; PSA 10 – 20ng/mL (=IR); or Gleason score≥8 or cancer core length≥4mm; PSA>20 ng/mL (=HR)	Sensitivity to detect CNS & CS PCa for SBx, and: .adjusted by relative sensitivity of TBx vs SBx, if TBx) .adjusted by relative sensitivity of 1st vs subsequent biopsy if 2nd biopsy	Decision Tree: classifies patients according to diagnostic accuracy, true disease status & underlying risk category + Markov model capturing treatment allocation conditional on classification & longer-term outcomes Health states: No PCa; unDx LR; unDx IR; unDx HR; unDx metastatic; Dx LR; Dx IR; Dx HR; Dx metastatic; PCa death; other cause death.	Via Markov model capturing disease onset and sequential disease progression from lower to higher risk category (LR→IR→HR) of localised disease and from HR to metastatic disease. PCa mortality only applies to metastatic disease.

Faria (2018)[135]/ Brown (2018)[136], UK To assess the CE of combinations of mpMRI, TRUS, TPMB for the diagnosis of PCa in men referred to 2ary care investigations	383 strategies with alternative combinations of mpMRI, TRUS, & TPMB, which differ in terms of: . whether or not, and when (to guide TRUS or to inform repeat biopsy) to use mpMRI; . the type of biopsy (TRUS-guided or TPM); .whether repeat biopsy is allowed and who receives it conditional on previous test results; . definition of suspicious lesion on mpMRI (4 alternative cut-offs) .definitions of CS PCa (2 alternatives)	Histopathological definition: 1. dominant Gleason pattern≥4 and/or any Gleason pattern 5 and/or cancer core length ≥6mm; or 2. any Gleason pattern≥4 and/or cancer core length≥4mm Clinical definition: - CNS PCa: PSA≥10ng/ml and Gleason score≥6 (=LR) - CS PCa: PSA 10-15 ng/ml & Gleason score (=IR); or Gleason score≥8 (=HR)	Probability of detecting PCa, CNS or CS PCa conditional on true risk category of LR, IR, HR) Specificity of detecting PCa	Decision Tree: classifies patients according to diagnostic accuracy, true disease status & underlying risk category + Markov model capturing treatment allocation & longer-term outcomes Health states: no PCa(?), localised PCa, metastatic disease, death	Via Markov model capturing disease progression from localised disease to metastatic disease. PCa mortality only applies to metastatic disease.
Barnett (2018)[134], US To assess the CE of diagnostic strategies involving MRI & TBx (alone or combined) for men undergoing biennial PSA screening	1. No PSA screening Screening strategies If PSA >4ng/mL: 2. SBx 3. MRI → PI-RADS<3: SBx; PI-RADS 3+: TBx 4. MRI → PI-RADS<3: no biopsy; PI-RADS 3+: TBx 5. mpMRI → PI-RADS<3: SBx; PI-RADS 3+: Combined biopsy 6. mpMRI → PI-RADS<3: no biopsy; PI-RADS 3+: Combined biopsy TBx performed with MRI fusion Individuals with biopsy (+) receive treatment, and those with biopsy (-) return to screening (not explicit)	Histopathological definition: . high-volume tumour and Gleason score 3+4 or Gleason score≥ 4+3 (high grade disease) Clinical definition: any Gleason score≥7	SBx: - Sensitivity of detecting PCa - Probability of incorrect grading for (+) biopsy TBx and combined biopsy: - sensitivity and specificity for high-grade cancer	Partially observable Markov model capturing screening/diagnostic outcomes (via implicit decision tree** embedded in the model), treatment allocation & longer-term outcomes Health states: .no PCa; other cause death; .pretreatment PCa states (unobservable): organ confined Gleason score<7, organ confined Gleason score=7, organ confined Gleason score>7, EPLN . detected PCa: PCa treatment (AS or RP), no recurrence following treatment, possible recurrence following treatment, metastatic PCa, PCa death.	Via partially observable Markov model capturing: . Onset of PCa . 1ary treatment allocation . Disease progression from localised to metastatic disease PCa mortality only applies to metastatic disease in detected states.

Patel (2018)[137], Netherlands To assess the CE of AS strategies for men with LR	1. 3-yearly SBx biopsy → biopsy (-): AS; biopsy (+): treated 2. 3-yearly mpMRI → mpMRI (-): AS; mpMRI (+): TBx → biopsy (-): AS; biopsy (+): treated 3. 3-yearly mpMRI → mpMRI (-): AS; mpMRI (+): treated mpMRI (+)/(-) defined in relation to presence of HR. All biopsies are performed via TRUS	Histopathological definition: Gleason score≥7 Clinical definition: NR, but text suggests that LR (PSA<10 ng/ml, Gleason score <6, and stage T2a) is equivalent to CNS PCa and HR (Gleason score≥7) to CS PCa	Sensitivity and specificity of detecting HR	Markov model capturing diagnostic outcomes (via implicit decision tree embedded in the Markov model) & longer-term outcomes Health states: LR, HR, survival after treatment LR, survival after treatment HR, death (due to PCa or other causes)	Via Markov model capturing disease progression from LR to HR. PCa mortality only applies to individuals with HR
Sathianathen (2018)[138], US To assess the CE of biomarkers in determining the need for biopsy in men with elevated PSA	1. SBx 2-5. biomarker → (< cut-off): followed-up (not explicit); (≥ cut-off): SBx 6. mpMRI → mpMRI(-): followed-up; mpMRI(+): TBx Biomarkers: phi, 4Kscore®, SelectMDx™ and the EPI (ExoDx™ Prostate [Intelli-Score]) Where individuals with biopsy (+) receive treatment, and those with biopsy (-) are followed-up. SBx is performed via TRUS. mpMRI (+)/(-) is not defined	NR	Sensitivity of detecting LG and HG PCa	Decision Tree: classifies patients according to diagnostic accuracy, true disease status + Markov model for Dx PCa (not described) + State transition model (not described) for unDx PCa capturing risk of clinical diagnoses due to symptoms and risk of metastasis by clinical diagnosis Health states: NR	NR
Pahwa (2017)[139], US To assess the CE of SBx & TBx (with alternative MRI-influence method (MRI fusion, cognitive fusion or in-bore) for biopsy-naïve men with elevated PSA &/or CS DRE	1. SBx 2-4. mpMRI→(no suspicious lesions): discharged; (suspicious lesions): TBx 5-7. mpMRI→(no suspicious lesions): SBx; (suspicious lesions): TBx Where individuals with biopsy (+) receive treatment, and those with biopsy (-) are discharged. All biopsies are performed via TRUS	Histopathological definition(?): CNS PCa: Gleason score<6 & tumour volume<0.5cm^3 Clinical definition: NR, but text suggests that CNS PCa is equivalent to LR and CS PCa to HR	Sensitivity for detecting PCa, CNS and CS PCa Specificity for PCa Probability of correctly classifying tumour aggressiveness	Decision tree classifies patients according to diagnostic accuracy, true disease status and allocates 1^{ary} treatment	Via lifetime health and cost payoffs conditional on diagnostic status (diagnosed/missed), 1^{ary} treatment, and age Pay-offs are informed by outcomes of an external Markov model (supplemented with assumptions for patient management options not examined in the external model)

Venderink (2017)[140], Netherlands To assess the CE of SBx & TBx (with alternative MRI-influence method used (MRI fusion or in-bore) for biopsy-naïve men with elevated PSA &/or abnormal DRE	1. SBx 2-3. mpMRI → (no suspicious lesions): discharged; (suspicious lesions): TBx (2. MRI fusion & 3. In-bore) Where individuals with biopsy (+) receive treatment, and those with biopsy (-) are discharged. All biopsies are performed via TRUS	Histopathological definition(?): Gleason score ≥3+4 high-volume (IR/ HR).	Sensitivity to detect CNS & CS PCa Specificity for PCa Probability of false CNS and CS	Decision Tree: classifies patients according to diagnostic accuracy, true disease status (CS & CNS PCa) & treatment allocation + Markov model capturing longer-term outcomes Health states: No PCa, status after RP, status after RT, status after AS, death	Via Markov model capturing long-term outcomes
Cerantola (2016)[141], Canada To assess the CE of using MRI and TBx for biopsy-naïve men with elevated PSA & abnormal DRE	1. SBx 2. mpMRI → (PI-RADS<3): followed-up; (PI-RADS≥3): TBx Where individuals with biopsy (+) receive treatment, and those with biopsy (-) are followed-up. SBx is performed via TRUS	NR CS PCa is not defined but manuscript suggests that it is equivalent to IR/HR	TBx: - Rate of biopsy(+) - Rate of CS among biopsy(+) SBx: - Rate of biopsy(+) - Rate of FN - Rate of CS among biopsy(+)	Markov model capturing diagnostic (via implicit decision tree embedded in the Markov model) & longer-term outcomes Health states: two set of health states 1. mpMRI, TBx, 2. SBx, SBx(+), (1) or (2) plus follow-up, LR PCa, IR/HR PCa, AS, curative treatment, biochemical recurrence, CRPC, PCa death, other-cause death	Via Markov model capturing: 1. biopsy alternatives: TBx or SB(+) 2. biopsy outcomes: No PCa (captured in follow-up), LR PCa, HR PCa; 3. 1^{ary} treatment allocation; 4. disease progression from localised disease (LR, IR/HR to relapse) to metastasis (CRPC) PCa mortality only applies to metastatic disease.
de Rooij (2014)[142], The Netherlands To assess the CE of using MRI and TBx for biopsy-naïve men with elevated PSA	1. SBx 2. mpMRI → (no suspicious lesions): followed-up; (suspicious lesions): TBx Where individuals with biopsy (+) receive treatment, and those with biopsy (-) are followed-up. SBx is performed via TRUS.	Histopathological definition: CNS PCa: Gleason score ≥3+4 or large tumour with Gleason score 3+3	Sensitivity and specificity for detecting PCa Probability of correctly classifying tumour aggressiveness	Decision Tree: classifies patients according to diagnostic accuracy, true disease status (CS & CNS PCa) & treatment allocation + Markov model capturing longer-term outcomes Health state: alive, dead	Via Markov model capturing long-term outcomes

Mowatt (2013)[143], UK To assess the CE of using alternative MRS/MRI sequences to target TRUS biopsy, compared with SBx in individuals with suspected PCa & a previous (-) biopsy	1. SBx (extended [14-16] core TRUS) 2. MRI/MRS → MRI/MRS (-): followed-up; MRI/MRS (+): TBx 3. MRI/MRS → MRI/MRS (-): SBx; MRI/MRS (+): TBx Individuals with biopsy (+) receive treatment, & those with biopsy (-) follow-up (with a repeat saturation at 12 months if FN)	NA	Sensitivity and specificity of detecting PCa	Markov model capturing diagnostic (via implicit decision tree embedded in the Markov model) & longer-term outcomes Health states: No PCa or undetectable PCa, Dx localised T1-2 PCa (LR), Dx localised PCa (IR), Dx localised PCa (HR), Dx locally advanced T3 PCa (or extraprostatic cancer), unDx localised T1-2 PCa (LR), unDx localised PCa (IR), unDx localised PCa (HR), unDx locally advanced T3 PCa, Dx metastatic PCa, PCa death, other cause death	Via Markov model capturing PCa onset, and disease progression from i) localised to metastatic PCa, and ii) from locally advanced to metastatic PCa. PCa mortality only applies to individuals with metastatic cancer.

*not explicit but there is some suggestion that low grade is referred to clinically non-significant and high grade to clinically significant; **decision rule in the original paper; (+) positive result; (-) negative result; ADT, androgen deprivation therapy; AS, active surveillance; CE, cost-effectiveness; CNS, clinically non-significant; CRPC, castration-resistant prostate cancer; CS, clinically significant; DRE, digital rectal examination; Dx, diagnosed; EBRT, external beam radiation therapy; FN, false negative;; GATP, general anaesthesia transperineal; GG, Grade Group; GS, Gleason score; HG, high grade prostate cancer); HR, high-risk prostate cancer; IR, intermediate-risk prostate cancer; ISUP, International Society of Urological Pathology; LATP, local anaesthesia transperineal; LATRUS, local anaesthesia transrectal ultrasound; LG, low grade prostate cancer; LR, low-risk prostate cancer; MRI, magnetic resonance image; MRS, magnetic resonance spectroscopy; NA, not applicable; NR, not reported; PCa, prostate cancer; PCA3, prostate cancer antigen 3; phi, prostate health index; PI-RADS, Prostate Imaging-Reporting and Data System; PSA, prostate specific antigen; RP, radical prostatectomy; RR, relative risk; RT, radiotherapy; RTx, radical treatment; SBx, systematic biopsy; T, tumour stage; TBx, targeted biopsy; TPMB, template prostate mapping biopsy; TRUS, transrectal ultrasound; w/wo, with or without; wo, without; WW, watchful waiting.

The majority of identified studies aimed to evaluate the cost-effectiveness of strategies for initial prostate cancer diagnosis involving biopsy approaches.[126, 129, 131, 133, 135, 136, 139-142] The study populations in some of these diagnostic studies included only biopsy-naïve individuals[131, 135, 136, 139-142] while others included biopsy- naïve individuals and those with a previous negative biopsy.[126, 133]The population in Mowatt et al. (2013)[143] included only individuals who had had a previous negative biopsy. Three studies evaluated alternative prostate diagnostic strategies in the context of PSA based screening.[130, 132, 134] One study examined alternative protocols of active surveillance for those diagnosed with low-risk prostate cancer.[137] One study examined the use of mpMRI and MRI-influenced biopsy as an alternative in the evaluation of prostate cancer biomarkers.[138]

In the majority of the identified studies a cohort simulation modelling approach using a combined decision tree and Markov model structure was applied.[126, 129, 131, 133, 135/136-140, 142, 143] In these models the decision tree component modelled the diagnostic/screening pathway to classify individuals according to their true disease and diagnostic outcomes, while the Markov model component linked the diagnostic outcomes (and subsequent clinical management decisions) to the long-term effects on outcomes. Other cohort models relied solely on a decision tree structure[139] or a Markov model structure[141] to evaluate the cost-effectiveness of the alternative strategies. One cohort model was described as a partially observable Markov model,[134] and distinguishes between unobservable pretreatment (or preclinical [i.e., prior to the presentation of any disease signs or symptoms]) and observable (or clinical) states. Two studies used a continuous-time microsimulation (i.e., patient level) model calibrated to registry data. These models had two main model components to reflect i) prostate cancer natural history (from preclinical to clinical cancer) and ii) its diagnostic and treatment pathways.[130, 132]

The biopsy diagnostic outcomes applied across studies allow classification of patients according to the presence of prostate cancer alone (i.e., no prostate cancer/prostate cancer),[143] or based on disease presence and its clinical significance (i.e., clinically nonsignificant/significant prostate cancer)[137, 139, 140]. One study classified patients according to ISUP grades into three categories: no prostate cancer (ISUP grade 0), prostate cancer of ISUP grade 1, or 2 and higher.[130] It is worth noting, that biopsy results only provide histopathological information, usually expressed in terms Gleason score and/or pattern (or as ISUP grade) and maximum core length. However, ascertaining the disease clinical significance for the purposes of guiding patient management requires knowledge of further prognostic information (e.g., T stage and PSA levels), as more radical treatment is only indicated for cancer with worse prognosis (i.e., those likely to progress at a quicker rate from localised to metastatic disease). The definition of clinical significance applied in the models to classify individuals according to biopsy results is based on the histopathological definition of clinical significance only. The full clinical definition of disease significance which is applied in the models to select patient management

is conditional on biopsy results and other prognostic information. Establishing a link between histopathological and clinical definitions of disease significance (usually requiring judgements on how to map across definitions and/or risk stratification) is thus a feature of most models. However, not all studies make a clear distinction between the two types of definitions of clinical significance.[131, 137, 139, 140, 142] In some studies, the definition of clinical significance is not provided.[132, 138, 141]

Some studies[126, 129, 131, 133, 135, 136] further classify prostate cancer according to three-tier cancer risk classifications, which are generally similar (generally low-risk, intermediate-risk, high-risk prostate cancer) despite some minor differences across classification in how each category is defined. While the exact definition of the risk categories varies across studies, individuals with prostate cancer are in general assigned to a risk category on the basis of their PSA levels, histopathological presentation and disease (T) stage.

In the majority of identified studies, the link between diagnostic outcomes and subsequent treatment choice was established via a Markov or partially observable Markov model component.[126, 129, 131, 133-143] The structure of most of these models allows capturing disease progression to metastatic disease[126, 129, 131, 133-137, 141] or high-risk disease.[137] The model by Barnett et al. (2018), allowed for progression in patients with undetected prostate cancer (preclinical states) across health states defined by Gleason score and whether disease localised, and from any of these states to metastatic disease. For patients with detected prostate cancer who underwent radical prostatectomy the progression to metastatic disease was done via a cancer recurrence health state.[134] In all of these disease progression models, prostate cancer mortality only applies to individuals with metastatic disease[126, 129, 131, 133-137, 141, 143] or high-risk disease.[137] Two Markov models did not consider disease progression, with long-term outcomes directly conditioned on true disease status, diagnostic status (diagnosed or undiagnosed cancer) and primary treatment received.[140, 142]

Hao et al. (2021)[130] and Getaneh et al. (2021)[132] modelled disease progression (and onset) within a calibrated microsimulation model. In Hao et al. (2021)[130] disease progression occurred sequentially from disease stage T1-T2 to T3-T4 and from T3-T4 to metastatic disease in preclinical states and from localised to metastatic disease in clinical states. Prostate cancer mortality only applied to individuals with metastatic disease in clinical states. In Getaneh et al. (2021)[132] disease onset was assumed to imply a T1 tumour stage; disease progression would occur sequentially from the T1 stage to T2, and from this to T3. At each tumour stage, individuals also progressed across Gleason scores (lower than 7 → equal to 7 → greater than 7). Individuals in each preclinical state could progress from local-regional to distant metastasis, but prostate cancer mortality only applied to individuals in clinical states.

In one study, long-term outcomes were quantified by the decision tree alone, which assigned lifetime QALY and cost pay-offs to each terminal node, conditional on true disease status, diagnostic status (diagnosed or missed) and allocated treatment.[139]

Of the sixteen studies identified at the first stage of the review, nine were selected for a more in-depth review, as these were identified as the most appropriate to support the conceptualisation of the *de novo* model given the relevance of:

- The comparisons and position in the diagnostic pathway –studies which compared biopsies conducted with MRI-influence methods (i.e., targeted and/or combined biopsies) for prostate cancer diagnosis;[129, 130, 134, 139, 140]
- UK policy relevance.[126, 131, 133, 135, 136]

Although Mowatt et al., 2013[143] was considered to have UK policy relevance, it was not considered for the second stage of this review, given that diagnostic accuracy in this study only allowed classifying individuals according to prostate cancer presence. Therefore, the evidence linkage in this study is unlikely to be suitable for the current decision problem, as the choice of prostate cancer management needs to be linked as a minimum to some level of prognostic information (e.g., clinical significance of disease).

Studies included in the model conceptualisation review

Table 33 in main text (see Section 116summarises the subset of identified studies included in the model conceptualisation review. A detailed description is provided next.

Population

The population in the majority of studies comprises individuals with suspected prostate cancer who enter a secondary care diagnostic pathway[126, 129, 131, 133, 135, 136, 139, 140], while other studies consider patients being screened for prostate cancer.[130, 134]

Some of the studies on patients with suspected PCa consider a single homogeneous population in terms of disease (and clinically significant disease) prevalence,[139, 140] others model different baseline populations defined by their diagnostic story (MRI results, number of previous biopsies)[126, 133] and underlying cancer risk category.[126, 133, 135, 136] One study further considers subgroups defined by age brackets, with increased disease prevalence for older individuals (but the same clinically significant prevalence for all subgroups).[139]

Hao et al. 2021[130] considered a population eligible for PSA-based prostate cancer screening. The manuscript mentions that individual heterogeneity is considered in the natural history model (informed by Swedish registry data) but does not clearly state which individual characteristics are modelled beyond PSA levels.

Biopsy approaches

A variety of biopsy approaches were compared in the studies; these differ by route of access (transrectal vs. transperineal), type of anaesthesia used (general vs. local anaesthesia), sample collection method (targeted vs. systematic vs. mapping or saturation biopsy) and MRI-influenced methods (software fusion, cognitive fusion, and in-bore MRI).

In the studies which compared alternative MRI-influenced methods with each other, one compared MRI followed by targeted biopsy approaches for those who tested positive on imaging with i) all three[139] or ii) just two methods (in-bore and software fusion)[140] vs. systematic biopsy (without prior MRI) for all patients. None of these studies specified the software fusion system modelled.

The study by Cheng et al., 2021[129] evaluated sequences of prostate biopsies with alternative combinations of i) systematic, ii) template mapping, and iii) combined targeted and systematic biopsy. The MRI-influenced method used for the combined biopsies was not specified. Another study considered a wide number of diagnostic strategies for patients with suspected prostate cancer, which included systematic, targeted and template mapping biopsies.[135, 136] No MRI-influenced method was specified for the targeted biopsy approaches in either study.

Two other studies compared diagnostic strategies with an MRI-influenced component (targeted alone or combined with systematic biopsy) vs. systematic biopsy, but in the context of PSA based screening.[130, 134] One study[130] did not specify whether MRI-influenced biopsies were performed with software fusion, cognitive fusion or in-bore methods. In the other study[134] MRI-influenced biopsies were conducted with software fusion, but the technology used was not specified.

The type of anaesthesia under which biopsies are performed is only specified for the studies which focus their comparison on transperineal vs. transrectal biopsy approaches.[126, 131] One assumes local anaesthesia for all biopsied patients regardless of biopsy route of access,[131] while the other considers local anaesthesia for those biopsied via the transrectal route and either general or local anaesthesia for transperineal biopsy.[126]

Souto-Ribeiro et al., 2022,[126]a previous DAR by the Southampton EAG, established two main comparisons between biopsy approaches: i) local anaesthesia transperineal (LATP) biopsy (with any type of biopsy device) vs. local anaesthesia transrectal (LATRUS) biopsy and general anaesthesia transperineal (GATP) biopsy and ii) LATP with specific freehand devices vs. LATRUS and vs.

transperineal transrectal biopsy conducted with a grid and stepping device conducted under local or general anaesthetic.

The NICE CG131 model[133] evaluated alternative follow-up strategies of individuals with suspected prostate cancer and placed little emphasis on alternative biopsy approaches. The main analysis presented results only for strategies which used transrectal biopsy, although strategies with transperineal mapping biopsy were considered in extended analyses only.

Another feature of the biopsy approaches modelled is whether repeat biopsies were allowed, the number of subsequent biopsies modelled and who would receive these. In the studies which considered the possibility of repeat biopsies this has been modelled in the following ways:

- All patients with a no prostate cancer diagnosis at previous biopsy were assumed to receive repeat biopsy with a maximum of one repeat biopsy allowed in the model (assumption not justified). It is not clear whether the repeat biopsy would follow the same biopsy approach as the index biopsy for all strategies, as only one strategy is fully illustrated.[131]
- All patients with a no prostate cancer diagnosis at previous biopsy were assumed to receive a repeat biopsy, in the subset of strategies allowing repeat biopsy.[129] Strategies were defined in terms of the number of repeat biopsies allowed (up to a maximum of 2) and on the sample collection method (combined, systematic or saturation) conditional on the method of the previous biopsy in the testing sequence. Repeat biopsies were assumed to always follow a sample collection method different from the one in previous biopsies in the testing sequence.
- A proportion of patients with a no prostate cancer or clinically non-significant prostate cancer diagnosis receive one repeat biopsy with LATRUS (regardless of biopsy approach for the index biopsy).[126] The proportion of patients receiving a repeat biopsy was informed by the literature (single centre observational study comparing TRUS, LATP and GATP biopsy) for the biopsy naïve populations, and by assumptions for those with previous biopsies (a lower proportion of repeat biopsy was assumed for the latter population). While the proportion of repeat biopsies was assumed to be the same across biopsy approaches in the base-case analysis for LATP, GATP, LATRUS, this assumption was relaxed in scenario analysis where LATRUS was assumed to result in more repeat biopsies than the transperineal biopsy approaches (LATP and GATP).
- Repeat biopsy was allowed across most strategies but depending on the strategy the biopsy would be performed in those diagnosed at index biopsy with i) no cancer, ii) clinically non-significant cancer or both no cancer and clinically non-significant cancer. The type of biopsy approach (template mapping, systematic or targeted) would also vary across strategy, but no strategy allowed more than one repeat biopsy.[135, 136]

Some studies did not model the possibility of repeat biopsy.[139, 140] In other studies, the possibility of repeat biopsy was not modelled within the diagnostic component of the strategies, but repeat biopsies for individuals who returned to screening and were identified again for biopsy via screening.[130, 134] The NICE CG131 model also did not consider consecutive biopsies in the diagnostic strategies.[133] All individuals with a 'no cancer' biopsy result returned to follow-up, but individuals could receive more than one biopsy if they tested positive again to the screening tests in their follow-up protocol.

Classification

In most studies, the diagnostic accuracy of the biopsy procedure classifies individuals as not having prostate cancer or having non-clinically significant or clinically significant prostate cancer.[126, 129, 131, 133-136, 139, 140] The exception was the study by Hao et al. (2021), in which classification is done by ISUP grade.[130] Both types of classification are usually defined by histopathological features of the biopsied lesions (graded according to Gleason scores).

The specificity of biopsy to detect prostate cancer is assumed perfect across most models, so individuals without prostate cancer cannot be misclassified as having the disease. However, studies differ in terms of other types of misclassification allowed for patients tested with biopsy procedures. Misclassification types allowed in the studies via both the structure and the parameterisation of the diagnostic accuracy for the biopsy approach include:

- Individuals with prostate cancer of any clinical significance diagnosed as not having the disease;[126, 129, 131, 133-136, 139, 140]
- Individuals with clinically significant prostate cancer misclassified as non-clinically significant;[126, 129, 131, 133-136, 139, 140]
- Individuals with clinically non-significant prostate cancer misclassified as clinically significant;[134, 139, 140]

Choice of clinical management

Decisions on patient management at diagnosis could be determined by the biopsy diagnostic outcomes alone[135, 136, 139, 140] or with other factors also influencing treatment allocation.[126, 129-131, 133, 134]

In three models[135, 136, 139, 140] patient management was attributed according to individuals' classification in terms of disease presence and clinical significance of disease. This classification was established based on the diagnostic accuracy of the biopsy approaches.

Some models, tracked the individuals underlying cancer prognostic risk and used this information jointly with the diagnostic outcomes to allocate treatment. For example, the Southampton DAR model[126] allocated treatments based on disease presence, clinical significance of disease and underlying cancer risk distribution. In order to classify patients according to these factors, the model

stratified individuals with prostate cancer into three cancer risk categories (low, intermediate, and high-risk) according to the lesion's Gleason score, disease stage and PSA levels in separate diagnostic sub-decision trees for individuals in each risk category (plus a sub-decision tree for individuals without prostate cancer). Low-risk disease was assumed to correspond to clinically non-significant disease (as determined by the diagnostic accuracy – i.e., based on Gleason score alone), and intermediate and high-risk disease to clinically significant disease.

Disease spread at diagnosis (localised vs. metastatic) was also considered a factor for treatment allocation in some studies,[126, 130, 133] which assumed that a proportion of individuals in the baseline population would have metastatic disease and, if disease was detected, received treatment with chemotherapy and/or androgen depleting therapy.

One study considered age and PSA levels alongside Gleason score to determine prostate cancer treatment allocation.[134] Patients older than 80 years old diagnosed with prostate cancer of any clinical significance were treated with watchful waiting. Patients diagnosed with clinically significant cancer and PSA levels higher than 20ng/mL or Gleason score greater than 8 would undergo tests for staging purposes. It is not clear how treatment was then allocated conditional on the results of staging.

In the model by Cheng et al. (2021),[129] treatment allocation was determined by diagnosed disease clinical significance, age (with palliative care for those 75 years old or older) and cancer risk category. Although, the text suggests that the distribution of treatments varies by diagnosed risk category, it is unclear how this is done since the biopsy only classifies patients according to clinical significance.

In summary, for patients diagnosed with prostate cancer, the primary treatment allocation was conditional on:

v. Diagnosed clinical significance of disease, true cancer risk category and disease spread;[126, 133]
vi. diagnosed disease clinical significance;[135, 136, 139, 140]
vii. Gleason score, PSA level and age;[134]
viii. type of biopsy (targeted or systematic), cancer risk category and age.[129]

In one study, the mechanism of treatment allocation for patients with diagnosed with cancer was not clear, but it may have been conditioned by ISUP grade (established by the biopsy diagnosis accuracy), disease T stage and spread. The manuscript suggests that the treatment pathways were informed by Swedish registry data, but does not describe how this was done.[130]

A range of evidence sources were used to inform the distribution of treatments for diagnosed prostate cancer. Amongst these the following are relevant in the UK context:

- The Southampton DAR model[126] based treatment distribution by risk category on UK clinical guidance and observed treatment allocation from national audit data.[144]
- The NICE NG131 model[133] used observed primary treatment distributions by risk category from UK registry data.[122]
- The PROMIS trial[135, 136] assumed that treatment choice was guided by diagnosed disease clinical significance alone.

Individuals diagnosed as not having prostate cancer were discharged to follow-up,[131, 133, 135, 136] or returned to the screening schedule.[130, 134] One study,[126] conditioned the individuals' subsequent management after a no prostate cancer diagnosis on whether they had been misclassified (TN results led to discharge and FN results [patients with prostate cancer of any risk category] to routine PSA monitoring). This assumption was not justified and it is not clear how in clinical practice the two groups of individuals (TN and FN) would be distinguished so that distinct treatment decisions could be made for each group.

Outcomes

The evidence linkage approaches applied in the identified studies to connect patient classification and subsequent treatment choices with longer-term outcomes differed in whether prostate cancer progression was explicitly modelled as an intermediate outcome or not.

Only two studies did not model disease progression.[139, 140] Pahwa et al. (2017)[139] conditioned lifetime QALYs and cost payoffs on diagnostic status (i.e., whether cancer had been diagnosed or remained undiagnosed), underlying true disease status (no prostate cancer, clinically non-significant or clinically significant prostate cancer) and type of treatment received. The model applied a life-expectancy multiplier, to adjust payoffs according to alternative starting ages (scenario analysis). The lifetime pay-offs were mainly derived from an external Markov model[128] comparing alternative treatments for patients with low-risk localised prostate cancer. The long-term Markov model in Venderink et al. (2017)[140] only allowed for transitions from alive to death states. Individuals with prostate cancer health states were defined in terms of the primary treatment received (status after i) active surveillance, ii) radical prostatectomy or iii) radiotherapy) or no treatment (for those who had been misclassified as not having cancer). In these patients, survival was conditional on type of treatment received and the underlying true disease clinical significance, with the diagnostic status (diagnosed vs. undiagnosed cancer) determining whether individuals received treatment.[140] In both these models, treatment had a direct impact on survival.[139, 140]

All other models considered disease progression from localised to metastatic disease although health states and possible state transitions varied across models.[126, 129, 131, 133-136] Some studies modelled progression from localised to metastatic disease, and conditioned disease progression on underlying

risk category and being correctly diagnosed/treatment received.[129, 131, 135, 136] Other studies modelled sequential disease progression across disease risk categories (from low to intermediate-risk and from the latter to high-risk disease) for localised disease followed by progression from the high-risk localised to metastatic disease. In these models, the probabilities of transitioning to later disease stages were conditioned on the underlying true disease status (including risk category) and being diagnosed as having clinically significant or non-significant disease.[126, 133] The screening studies modelled progression differently in the preclinical and clinical states.[130] In the microsimulation model,[130] individuals with prostate cancer could transition between preclinical states defined in terms of ISUP grade, tumour stage and metastasis; within each ISUP grade individuals progressed sequentially from stage T1-T2 to T3-T4 and from T3-T4 to metastatic disease. In the clinical states (for those whose prostate cancer was detected) disease progression occurred from localised to metastatic disease. In the partially observed Markov model,[134] disease progression in the preclinical states could occur i) sequentially between three localised disease health states defined according to Gleason score (<7, =7, >7) ii) from any of the localised disease states to extra-prostatic or lymph node-positive cancer, or iii) from any of the preclinical states to observable (clinical) metastatic cancer. The rate of progression to metastatic cancer was the same for all pre-clinical states. In the clinical states, patients treated with radical prostatectomy could transition to one of the two post-treatment states: no recurrence following treatment (NRFT) or possible recurrence following treatment (PRFT) health states. Progression to metastatic cancer was only possible for individuals in the PRFT state, with those in the NRFT state assumed cured. The probability of transitioning from the prostate cancer treatment health state to the post-treatment states was conditional on disease location (organ confined vs. extra prostatic or lymph node-positive cancer) and treatment received. Patients who were treated with active surveillance could progress to metastatic disease at the same rate as those who were untreated, unless they transitioned to surgical treatment. The model appears to track progression over time across Gleason scores and disease location for those under active surveillance, in a manner similar to what happened in the pre-clinical states.

All the disease progression models shared the assumption that prostate cancer mortality only applied to patients with metastatic disease. Treatment for patients identified as having cancer reduced disease progression to metastatic cancer compared to untreated patients, and thus reduced the probability of dying from prostate cancer for these patients. The transition probabilities for treated and untreated patients in the Markov disease progression were estimated by calibration or partially observable Markov model decision processes (as progression is an unobservable process). The data sources and calibration methods used to estimate these transition probabilities differed across models, and are reviewed below.

The PROMIS model[135, 136] calibrated the probability of progressing from localised to metastatic disease by risk category and treatment received, combining risk stratified survival data and proportion of patients with metastases from the Prostate Cancer Intervention versus Observation Trial (PIVOT),[119] with the mortality in the metastatic subgroup of the Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy (STAMPEDE) trial.[123] The PIVOT observation arm was used to inform the transition probabilities for individuals with prostate cancer who did not receive active treatment (due to correct classification on misclassification depending on the risk category). The PIVOT radical prostatectomy arm was used to inform the transition probabilities for those treated with active treatment (true positives with intermediate and high-risk cancer). The 'treatment' effects of being diagnosed on disease progression were thus informed by randomised comparative efficacy evidence.

The models which disaggregated disease progression by cancer risk categories, also used calibration to estimate transition probabilities.[126, 133] The calibration method estimated transition probabilities first for the transition from high-risk to metastatic disease, then from intermediate to high-risk disease, and finally from low-risk to intermediate-risk disease can be derived. The calibration was done separately for the undetected and detected cancers using different data sources. Transition probabilities for the undetected cancers used cumulative metastases risk rates by cancer risk category from the watchful waiting arm in the Scandinavian Prostate Cancer Group Study Number 4 (SPCG4) trial[145] jointly with and Swedish life-table data (from 1999 to reflect background mortality in the trial). For the diagnosed cancers, the data sources for calibration included: cancer-specific survival by risk category sourced from a UK registry study,[122] all-cause survival for people with metastatic prostate cancer from the STAMPEDE trial,[59] and UK life-table (from 2010-2012 to reflect background trial mortality in STAMPEDE). Thus, this calibration approach relies on an indirect naïve comparison to derive the 'treatment' effects of being diagnosed on disease progression, which may introduce bias on the probabilities of disease progression used in the model.

The screening model by Hao et al. 2021[130] does not describe the calibration method used to parameterise disease progression transitions, mentioning only that the model is calibrated to UK registry data. The other screening model[134] also used calibration to estimate the transition probability from localised and extra-prostatic or lymph node-positive cancer (preclinical states). The authors varied the metastasis rate in 10-year periods and calibrated the values so that the resulting age-dependent risk of prostate cancer specific death under routine screening matched the values estimated from historical US cancer registry data. For the clinical states, the authors state that the probability of transitioning from recurrent to metastatic disease was informed by another US cancer registry data and using the methodology of an external partially observed Markov model. It is not clear how the methodology described for the external model was applied in the model developed by Barnett et al.,

2018. It is also not clear why the transition probabilities for the preclinical and clinical states were estimated by two different methods (i.e., calibration and partially observed Markov decision process).

One model[129] does not describe how transition probabilities were estimated and does not fully report the data sources used to inform these parameters.

In general disease progression models, survival outcomes for individuals with prostate cancer were conditional on having metastatic disease and age. Two models[126, 133] further conditioned mortality on whether metastatic disease was diagnosed (and therefore, received treatment for metastatic cancer) or not. Metastatic mortality data sources of relevance to the UK context include different publications of the STAMPEDE study, a UK based trial which compared the survival outcomes of men with newly diagnosed metastatic, high risk or node positive cancer treated with alternative cancer treatments. The PROMIS and related models estimated the probability of metastatic death using early (median follow-up of 20 months) survival data of men with newly diagnosed metastatic prostate cancer from the control arm (who received standard of care consisting of androgen depleting therapy) of the STAMPEDE trial. The NICE NG131 and related models, used a later survival data cut (median follow-up 43 months) from the docetaxel and control arms of the STAMPEDE trial that includes individuals with metastatic and non-metastatic disease.[59]

HRQoL outcomes of patients with prostate cancer were conditional on:

- Having metastatic disease[126, 129-131, 133-136] – negative impact on HRQoL;
- Having castration resistant metastatic disease[129] – negative impact on HRQoL;
- age[126, 129-131, 133-136] – decreasing utility with age;
- Being diagnosed with prostate cancer[130, 139, 140] – negative impact on HRQoL;
- Receiving radical treatment[129] – positive impact on HRQoL;
- Underlying true disease status (including clinical significance)[139] – negative impact on HRQoL of having prostate cancer, which is worsened by presence of clinically significant disease;
- Adverse events with radical treatment by true risk category[126, 133] – negative impact on HRQoL;
- Treatment received and time since treatment initiation[130, 134, 140] – initial negative impact on HRQoL with improvement in post-treatment period;
- End of life[130, 134] – negative impact on HRQoL.

The UK relevant utility sources for patients with prostate cancer in the long-term outcome models include:

- Torvinen et al., 2013[146] – for the disutility of metastatic disease
- Ara and Brazier, 2010[147] – for the disutility of ageing
- Mowatt et al., 2013[143] – for the disutility of treatment related adverse events (combined with rates of adverse events from Donovan et al., 2016[148])

The long-term HRQoL outcomes of patients without prostate cancer were dependent on age in most models,[129, 131, 133-136] with Ara and Brazier, 2010[147] the most frequently used source to inform age adjusted utilities.

Most models considered the cost of treatment for patients with diagnosed localised or locally advanced prostate cancer (radical treatment or active surveillance)[126, 129-131, 133-136, 139, 140] and management of treatment adverse events.[126, 131, 133, 135, 136] Patients with undiagnosed prostate cancer would incur the costs of routine follow-up[126, 129, 131, 133, 135, 136, 139] or of delayed radical treatment.[139] The studies also considered the costs of metastatic disease treatment with or without staging and follow-up tests.[126, 129, 131, 133-136] Two models assumed diagnosed metastatic disease would be treated differently if diagnosed (docetaxel would be added to androgen depleting therapy) compared to undiagnosed metastatic disease and that treatment with docetaxel would vary with age.[126, 133] Some models included an end-of-life cost for patients who died from prostate cancer,[126, 129, 130, 133, 134] with one study conditioning the end-of-life costs on age at death.[134]

The costs of individuals who did not have prostate cancer were not clearly reported for most models, but, where reported, consisted of the costs of routine follow-up.[126, 129, 133, 134]

In UK relevant models, treatment and follow-up resource use was informed mainly by UK (clinical and technology appraisal) guidance, as well as other published data (for example, a randomised control trial informed adverse event rates of treatment[148]) and supplemented with assumptions. End-of-life costs were uprated to the relevant price year based on Round et al. (2015)[149] Unit costs were sourced mainly from national published sources.

APPENDIX 10. EXTENSION OF THE EVIDENCE SYNTHESIS TO DETERMINE DIAGNOSTIC ACCURACY

Methods

Description of methods

Methods were developed to provide an internally consistent framework for evidence on the distribution of test results across a number of technologies (from the evidence synthesis), and data on the extent of misclassification of the technologies in relation to (true) disease status.

This framework relies on expressing the natural probability relationships between the different quantities of interest. The extent of misclassification is made explicit by the accuracy matrix. The accuracy matrix was expressed using conditional probabilities, with its elements being the probability of obtaining a particular test result with one method conditional on a particular level of (true) disease status, that is, the probability of a test (A) retrieving a particular result x in patients with a particular disease (D) level y – P[A=x|D=y] – or, using simplified notation, $p^{(A)}{}_{x|y\cdot}$. The set of conditional probabilities that fully define accuracy are shown in the matrix in Table 83. Together with prevalence estimates, P[D=y], or p_y in the simplified notation shown at the left side of Table 83, this matrix determines the distribution of test results, P[A=x], shown at the top of Table 83 using the simplified notation of $p^{(A)}{}_x$.

Table 83: Relationship between distribution of test results, the prevalence and the accuracy matrix

prevalence		distribution of test results					distribution of test results				
		$p^{(1)}{}_1$	$p^{(1)}{}_2$	$p^{(1)}{}_3$	$p^{(1)}{}_4$	$p^{(1)}{}_5$	$p^{(3)}{}_1$	$p^{(3)}{}_2$	$p^{(3)}{}_3$	$p^{(3)}{}_4$	$p^{(3)}{}_5$
		cognitive fusion, (1)					software fusion, (2)				
	D	1	2	3	4	5	1	2	3	4	5
p_1	1	$p^{(1)}{}_{1\|1}$	0	0	0	0	$p^{(3)}{}_{1\|1}$	0	0	0	0
p_2	2	$p^{(1)}{}_{1\|2}$	$p^{(1)}{}_{2\|2}$	0	0	0	$p^{(3)}{}_{1\|2}$	$p^{(3)}{}_{2\|2}$	0	0	0
p_3	3	$p^{(1)}{}_{1\|3}$	$p^{(1)}{}_{2\|3}$	$p^{(1)}{}_{3\|3}$	0	0	$p^{(3)}{}_{1\|3}$	$p^{(3)}{}_{2\|3}$	$p^{(3)}{}_{3\|3}$	0	0
p_4	4	$p^{(1)}{}_{1\|4}$	$p^{(1)}{}_{2\|4}$	$p^{(1)}{}_{3\|4}$	$p^{(1)}{}_{4\|4}$	0	$p^{(3)}{}_{1\|4}$	$p^{(3)}{}_{2\|4}$	$p^{(3)}{}_{3\|4}$	$p^{(3)}{}_{4\|4}$	0
p_5	5	$p^{(1)}{}_{1\|5}$	$p^{(1)}{}_{2\|5}$	$p^{(1)}{}_{3\|5}$	$p^{(1)}{}_{4\|5}$	$p^{(1)}{}_{5\|5}$	$p^{(3)}{}_{1\|5}$	$p^{(3)}{}_{2\|5}$	$p^{(3)}{}_{3\|5}$	$p^{(3)}{}_{4\|5}$	$p^{(3)}{}_{5\|5}$
		accuracy matrix					accuracy matrix				

Note that, due to the nature of biopsy and histological examination of the biopsy specimen, it is reasonable to assume that false positive results are not possible, i.e., if cancer is histologically identified, then it is present. This implies that biopsy methods cannot identify a higher ISUP Grade

than true disease status, and therefore zero probability is attributed to such cases in the above accuracy matrix.

Where two methods are of interest, the problem becomes more complex. Table 83 formalises the problem by depicting the quantities of interest for two alternative biopsy methods, including the prevalence (i.e., the true distribution across GG), which is independent of test results, the two conditional accuracy matrices, and the (marginal) distributions of test results, which are themselves a function of prevalence and accuracy. The key relationships that introduce complexity are:

- prevalence is independent of test results and therefore a common prevalence estimate needs to ground all distributions of test results, and be consistent with these, and
- explicit accounts of accuracy need to respect both the prevalence estimates and the marginal distribution estimates derived from the synthesis.

Where the distribution of test results has been related across tests without consideration for their accuracy against a reference standard, a structured approach is therefore required for characterising accuracy to ensure that probability relationships are maintained.

Note that such a model does not identify concordance between methods in biopsy test results. To consider concordance, the synthesis model would have had to be grounded on the underlying joint or conditional probabilities of classification across tests that, i.e., the likelihood of identifying individuals in a particular category using one method and in another category using a different method (joint probabilities) or the likelihood of individuals identified in a particular category by one method being classified in another by a different method (conditional probabilities). Joint/conditional probabilities determine the potential concordance between tests, which cannot be ascertained by the marginal distributions alone, that is, the same marginal distributions can be retrieved under very different levels of concordance between tests.

The approach developed for the current assessment was designed to:

- be grounded on the results of the evidence synthesis model,
- return a true distribution across Gleason Group categories (prevalence) that is internally valid, i.e., that is not lower than the estimated GG detection rates of the different biopsy methods,
- be grounded on available evidence on the likely accuracy of MRI fusion conditional on GG, and
- define accuracy matrices for the remaining biopsy methods of interest that are consistent with both prevalence and the distributions of biopsy results from the evidence synthesis.

To achieve this, an extension to the synthesis model was developed in WinBUGS,[150] drawing on the broader evidence in Section 4.2.2. To allow for an internally consistent approach, we grounded our methodology on evidence of the distribution of test results obtained with targeted-MRI methods, and of their accuracy. Given that disease prevalence is fully determined by these two results, the prevalence evidence identified in Section ## will not be explicitly incorporated in our analyses but will instead be used qualitatively to put our results into context.

1) Describing distribution of test results

The distributions of test results across the disease categories for the relevant biopsy methods within each disconnected network of Model 1a were computed by applying network-specific baseline distributions to the results of the network meta-analysis. Building from the analyses in the evidence synthesis section, the baseline distributions were assumed uncertain by using a Multinomial likelihood to describe the data from the empirical studies and an uninformative Dirichlet prior for its hyperparameters. The Dirichlet prior was implemented via a series of conditional Beta distributions to facilitate the later use of constraints.

Note that the scope of this assessment is to compare targeted biopsy methods; therefore, results on systematic biopsy used in isolation will not be shown here (by not including the broader literature on the accuracy of systematic biopsy, the results are also not relevant to support decision making).

2) Describing the accuracy matrix for software fusion

Evidence on the accuracy of software fusion in identifying disease status according to the categories of interest was used to characterise this probabilistically in the model. A Multinomial likelihood was used to describe the distribution of test results conditional on each particular level of true disease status (each line in the matrix in Table 83). The hyperparameters of the Multinomial were attributed an uninformative Dirichlet distribution, implemented via a series of conditional Beta distributions to facilitate the later use of constraints.

3) Deriving the prevalence distribution

The derivation of prevalence followed two steps.

3.1) Analytical derivation of prevalence from the marginal distribution and accuracy matrix for cognitive fusion

The prevalence and the accuracy matrix for a particular technology fully define the marginal distribution of test results for that technology. If represented in matrix form, the prevalence vector, $\mathbf{p}$, multiplied by the accuracy matrix, $\mathbf{M}$, retrieves the test result marginal distribution, $\mathbf{p}^{(i)}$, i.e., $\mathbf{p} \cdot \mathbf{M} =$

$\mathbf{p}^{(i)}$. We have used this relationship to derive the distribution of prevalence, i.e., $\mathbf{p} = \mathbf{p}^{(i)} / \mathbf{M}$. Because of the reverse calculation, a constraint was implemented to ensure prevalence results across categories would sum to 1.

3.2) Derivation and application of constraints for the prevalence distribution

Given the absence of false positive results (i.e., that biopsy cannot retrieve results of GG higher than the true value), the true distribution of disease across Gleason Groups is constrained by the marginal distributions of test results obtained across tests. This is because the prevalence of higher-grade tumours is expected to be at least equal to the maximum proportion in those groups identified across all tests. This means that:

- the true prevalence of GG 4 or 5 (j=5) is equal or higher than the maximum proportion of GG 4 or 5 identified across all tests – $p_5 \geq \max_i(p^{(i)}_5)$;
- the true prevalence of histology GG 3 and above (j=4 or j=5) is equal or higher than the maximum proportion of GG 3 and above identified across all tests – $p_5 + p_4 \geq \max_i(p^{(i)}_5 + p^{(i)}_4)$;
- the true prevalence of histology GG 2 and above (j=3, j=4 or j=5) is equal or higher than the maximum proportion of GG 2 and above identified across all tests – $p_5 + p_4 + p_3 \geq \max_i(p^{(i)}_5 + p^{(i)}_4 + p^{(i)}_3)$; and
- the true prevalence of histology GG 1 and above (j=2, j=3, j=4 or j=5) is equal or higher than the maximum proportion of GG 1 and above identified across all tests – $p_5 + p_4 + p_3 + p_2 \geq \max_i(p^{(i)}_5 + p^{(i)}_4 + p^{(i)}_3 + p^{(i)}_2)$.

The true prevalence distribution should meet these conditions. The boundaries for each of the inequalities defined (i.e., the values at equality) can be determined recursively (with calculations starting at the highest grade). These conditions were implemented in WinBUGS using inequality-constrains (see code below).

4) Derivation of accuracy matrix for other technologies

The accuracy matrix for the remaining technologies is determined by the prevalence estimates and by their marginal distributions. The diagonal cells in each of the accuracy matrices were therefore defined as a function of prevalence, probability of test result and other relevant elements in the accuracy matrices, by using the structural relationships between these parameters. For example, for category 4, the diagonal of the accuracy matrix for biopsy method k was defined as:

$p^{(k)}_{4|4} = (p^{(k)}_4 - p_5 \cdot p^{(k)}_{4|5}) / p_4$,

which subtracts those from category 5 that were incorrectly identified as 4's from the total with category 4 test results.

The remaining free elements of each line in the matrix were sampled from an uninformative Dirichlet distribution (defined as a set of conditional Beta distributions). Given that the diagonal cells relating prevalence with distribution of test results used the non-diagonal elements of the matrix, information is already conveyed on these parameters, and therefore final inference on these will not be fully uninformative. All accuracy parameters were constrained to be between 0 and 1, as the inverse matrix calculation, on its own, does not ensure that.

Implementation

The extension to the evidence synthesis model was developed in WinBUGS and was appended to the synthesis model code to draw on the inferences from the synthesised log odds ratios. The constraints implemented within the code extension need the log odds ratios in the synthesis model to be influenced by these. This will ensure that the inferences on the log odds ratio from the extended model are plausible with the data incorporated (accuracy matrices and baseline distribution of test results) and with the structural relationships between the quantities of interest. To evaluate the influence over the unconstrained evidence synthesis inferences, we will compare the probabilities of test results derived from the synthesis model used in isolation (Section 4.4.2.1) with those derived from the extended synthesis and accuracy model.

Additionally, non-diagonal elements of the accuracy matrices inferred by the model were simulated from a stochastic distribution, with information on them conveyed indirectly via the diagonal elements. For this reason, retrieving test results from inferences over the prevalence and accuracy matrix approximates, but does not equal, the distribution of test results retrieved by the synthesis model. Results were therefore also compared to determine the magnitude of differences.

WinBUGS code for extended synthesis model

Code

```
model{
for (i in 1:ns){ # studies reporting all categories 1,2,3,4,5
  for (k in 1:na[i]) {
    y[i,k,1:nc] ~ dmulti(q[i,k,1:nc], M[i,k])
    for (r in 1:nc) {
      q[i,k,r] <- phi[i,k,r]/sum(phi[i,k,])
      log(phi[i,k,r]) <- a[i,r] + d[t[i,k],r] - d[t[i,1],r]
     # predicted number events
     yhat[i,k,r] <-  q[i,k,r] * M[i,k]
     # Deviance contribution
     dv[i,k,r] <- 2*y[i,k,r]*(log(y[i,k,r]/yhat[i,k,r]))
     }
    dev[i,k] <- sum(dv[i,k,]) # deviance contribution of each arm
   }
```

```
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  for (r in 2:nc) {a[i,r] ~ dnorm(0, 0.0001)}
  a[i,1] <- 0
   # summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
 }
#relative effects of treatment 1 compared to itself are zero, for all categories
for (r in 2:nc) {d[1,r] <- 0}
for (k in 1:nt){
  # giving phi[i,k,1] = 1, logOR of  going from cat 1 to cat 1 for all treats
  d[k,1] <- 0
  for (r in 2:nc) {
    # vague priors for relative treatment effects: log-odds ratios
    d[k,r] ~ dnorm(0, 0.0001)
   }
 }
# STUDIES WITH COLLAPSED CATEGORIES: TYPE A
for (i in (ns+1):(ns+nsA)){ # studies reporting categories 1,2-5
  for (k in 1:na[i]) {
    y[i,k,1] ~ dbin(q[i,k,1], M[i,k])
    # first category the same
    q[i,k,1] <- phi[i,k,1]/sum(phi[i,k,1:ncA])
    log(phi[i,k,1]) <- a[i,1] + d[t[i,k],1] - d[t[i,1],1]
     yhat[i,k,1] <-  q[i,k,1] * M[i,k]
     # Deviance contribution
     dev[i,k] <- 2 * (y[i,k,1] * (log(y[i,k,1])-log(yhat[i,k,1])) +  (M[i,k]-
y[i,k,1]) * (log(M[i,k]-y[i,k,1]) - log(M[i,k]-yhat[i,k,1])))
# last category is collapsed, type A
    q[i,k,2] <- 1- q[i,k,1]
    log(phi[i,k,2]) <- a[i,2] + dA[t[i,k],2] - dA[t[i,1],2]
   }
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  a[i,2] ~ dnorm(0, 0.0001)
  a[i,1] <- 0
 # summed residual deviance contribution for this trial
 resdev[i] <- sum(dev[i,1:na[i]])
 }
dA[1,2] <- 0
for (k in 1:nt){
  # vague priors for relative treatment effects: log-odds ratios
  dA[k,2] ~ dnorm(0, 0.0001)
   }
# STUDIES WITH COLLAPSED CATEGORIES: TYPE B
for (i in (ns+nsA+1):(ns+nsA+nsB)){ # studies reporting categories 1,2,3-5
  for (k in 1:na[i]) {
    y[i,k,1:ncB] ~ dmulti(q[i,k,1:ncB], M[i,k])
    for (r in 1:(ncB-1)) {  # first 2 categories the same
      q[i,k,r] <- phi[i,k,r]/sum(phi[i,k,1:ncB])
      log(phi[i,k,r]) <- a[i,r] + d[t[i,k],r] - d[t[i,1],r]
     # predicted number events
     yhat[i,k,r] <-  q[i,k,r] * M[i,k]
     # Deviance contribution
     dv[i,k,r] <- 2*y[i,k,r]*(log(y[i,k,r]/yhat[i,k,r]))
     }
# last category is collapsed, type B
    q[i,k,3] <- 1- sum(q[i,k,1:(ncB-1)])
    log(phi[i,k,3]) <- a[i,3] + dB[t[i,k],3] - dB[t[i,1],3]
   # predicted number events
   yhat[i,k,3] <-  q[i,k,3] * M[i,k]
   # Deviance contribution
    dv[i,k,3] <- 2*y[i,k,3]*(log(y[i,k,3]/yhat[i,k,3]))
     dev[i,k] <- sum(dv[i,k,1:ncB]) # deviance contribution of each arm
   }
```

```
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  for (r in 2:ncB) {a[i,r] ~ dnorm(0, 0.0001)}
  a[i,1] <- 0
   # summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
 }
dB[1,ncB] <- 0
for (k in 1:nt){
  # vague priors for relative treatment effects: log-odds ratios
  dB[k,ncB] ~ dnorm(0, 0.0001)
   }
# STUDIES WITH COLLAPSED CATEGORIES: TYPE C
for (i in (ns+nsA+nsB+1):(ns+nsA+nsB+nsC)){ # studies reporting categories 1,2,3,4-
5
  for (k in 1:na[i]) {
    y[i,k,1:ncC] ~ dmulti(q[i,k,1:ncC], M[i,k])
    for (r in 1:(ncC-1)) {  # first 3 categories the same
      q[i,k,r] <- phi[i,k,r]/sum(phi[i,k,1:ncC])
      log(phi[i,k,r]) <- a[i,r] + d[t[i,k],r] - d[t[i,1],r]
      # predicted number events
      yhat[i,k,r] <-  q[i,k,r] * M[i,k]
      # Deviance contribution
      dv[i,k,r] <- 2*y[i,k,r]*(log(y[i,k,r]/yhat[i,k,r]))
     }
# last category is collapsed, type C
    q[i,k,4] <- 1- sum(q[i,k,1:(ncC-1)])
    log(phi[i,k,4]) <- a[i,4] + dC[t[i,k],4] - dC[t[i,1],4]
   # predicted number events
   yhat[i,k,4] <-  q[i,k,4] * M[i,k]
   # Deviance contribution
   dv[i,k,4] <- 2*y[i,k,4]*(log(y[i,k,4]/yhat[i,k,4]))
   dev[i,k] <- sum(dv[i,k,1:ncC]) # deviance contribution of each arm
   }
  # vague priors for BL log odds of transition from 1st category to cat r in study
i
  for (r in 2:ncC) {a[i,r] ~ dnorm(0, 0.0001)}
  a[i,1] <- 0
   # summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
 }
dC[1,ncC] <- 0
for (k in 1:nt){
  # vague priors for relative treatment effects: log-odds ratios
  dC[k,ncC] ~ dnorm(0, 0.0001)
   }
totresdev <- sum(resdev[])        # Total Residual Deviance
# pairwise ORs and LORs for all possible pair-wise comparisons
for (c in 1:(nt-1)) {
   for (k in (c+1):nt)  {
      for (r in 1:nc) {
        or[c,k,r] <- exp(d[k,r] - d[c,r])
        lor[c,k,r] <- (d[k,r]-d[c,r])
         }
     }
 }
# calculate absolute probabilities from relative effects (no uncertainty in
baseline)
# baseline intervention = 3 (software)
for (r in 1:nc){
  T[3,r] <- b1[r]  # baseline probabilities for software biopsy (from data)
  T[5,r] <- b2[r]  # baseline probabilities for software + SB (from data)
  # log-odds of being classified in category r using intervention 3=software
  A.T[r] <- log(T[3,r]/T[3,1])
  # log-odds of being classified in category r using intervention 5=software +SB
  B.T[r] <- log(T[5,r]/T[5,1])
```

```
 }
for (k in 1:2){                  # fully connected only T[2,] to T[3,]
  for (r in 1:nc){
    phi.T[k,r]  <- exp(A.T[r] - lor[k,3,r])
    T[k,r] <- phi.T[k,r]/(sum(phi.T[k,]))
   }
 }
for (r in 1:nc){
  phi.T[4,r]  <- exp(B.T[r] - lor[4,5,r])
  T[4,r] <- phi.T[4,r]/(sum(phi.T[4,]))
 }


# EXTENSION
####### calculation of probabilities  #######

##### baseline likelihood for ref biopsy method
## multinomial with Dirichlet vague prior implemented as conditional Betas (to
allow constraints ahead)

### baseline likelihood in 1st network, p[3,]
B.y[1:nc] ~ dmulti(p[3,1:nc], B.M)
for (i in 1:nc){ ax[i] <- 1 } # define parameters of Dirichlet distribution
p[3,1] ~ dbeta(ax[1], bx[1])
bx[1] <- nc-1
for (i in 2:(nc-1)) {
       aux[i] ~ dbeta(ax[i],bx[i])
       bx[i] <- sum(ax[(i+1):nc])
  p[3,i] <- (1 - sum(p[3,1:(i-1)])) * aux[i]
 }
p[3,nc] <- 1 - sum(p[3,1:(nc-1)])

### baseline likelihood in 2nd network, p[4,]
B.z[1:nc] ~ dmulti(p[5,1:nc], B.Mz)
for (i in 1:nc){ ax1[i] <- 1 }
p[5,1] ~ dbeta(ax1[1], bx1[1])
bx1[1] <- nc-1
for (i in 2:(nc-1)) {
       aux1[i] ~ dbeta(ax1[i],bx1[i])
       bx1[i] <- sum(ax1[(i+1):nc])
  p[5,i] <- (1 - sum(p[5,1:(i-1)])) * aux1[i]
 }
p[5,nc] <- 1 - sum(p[5,1:(nc-1)])


### calculation of probabilities for both networks from relative effects, lor
for (r in 1:nc) {
             A.1[r] <- log(p[3,r]/p[3,1])
             A.2[r] <- log(p[5,r]/p[5,1])
             }
for (k in 1:2){
 for (r in 1:nc){
             phi.B[k,r]  <- exp(A.1[r] - lor[k,3,r])
             p[k,r] <- phi.B[k,r]/(sum(phi.B[k,]))
   }
 }
for (r in 1:nc){
  phi.B[4,r]  <- exp(A.2[r] - lor[4,5,r])
  p[4,r] <- phi.B[4,r]/(sum(phi.B[4,]))
}


####### determining prevalence   #######
#-- likelihood of conditional accuracy matrix based on external evidence -- FUS
# pac[true disease status, test result]

##GG1
pac[3,1,1] <- 1
```

```
##GG2
y2.ac[1:2] ~ dmulti(pac[3,2,1:2], M.ac[1])
pac[3,2,1] ~ dbeta(p2.ac[1], p2.ac[2])       # non-inf beta prior for 1st
probability
for (k in 1:2) { p2.ac[k] <- 1}
pac[3,2,2] <- 1- pac[3,2,1]

##GG3
      y3.ac[1:3] ~ dmulti(pac[3,3,1:3], M.ac[2])   # true status GG2 (category 3)
# Dirichtlet prior implemented as conditional beta distributions
for (i in 1:3){ p3.ac1[i] <- 1 } # define parameters of Dirichlet distr for this
category
# conditioning start at 1
pac[3,3,1] ~ dbeta(p3.ac1[1], p3.ac2[1])
p3.ac2[1] <- sum(p3.ac1[2:3])
p3.ac2[2] <- p3.ac1[3]
p3.aux ~ dbeta(p3.ac1[2],p3.ac2[2])
pac[3,3,2] <- (1 - pac[3,3,1]) * p3.aux
pac[3,3,3] <- 1 - sum(pac[3,3,1:2])

##GG4
y4.ac[1:4] ~ dmulti(pac[3,4,1:4], M.ac[3])  # true status GG3 (category 4)
# Dirichtlet prior implemented as conditional beta distributions
for (i in 1:4){ p4.ac1[i] <- 1 } # define parameters of Dirichlet distr for this
category
# conditioning start at 1
pac[3,4,1] ~ dbeta(p4.ac1[1], p4.ac2[1])
for (i in 1:3){ p4.ac2[i] <- sum(p4.ac1[(i+1):4])  }
for (i in 2:3) {
  p4.aux[i] ~ dbeta(p4.ac1[i],p4.ac2[i])
  pac[3,4,i] <- (1 - sum(pac[3,4,1:(i-1)])) * p4.aux[i]
 }
pac[3,4,4] <- 1 - sum(pac[3,4,1:3])

##GG5
y5.ac[1:5] ~ dmulti(pac[3,5,1:5], M.ac[4])  # true status GG4 or GG5 (category 5)
# Dirichtlet prior implemented as conditional beta distributions
for (i in 1:5){ p5.ac1[i] <- 1 } # define parameters of Dirichlet distr for this
category
# conditioning start at 1
pac[3,5,1] ~ dbeta(p5.ac1[1], p5.ac2[1])
for (i in 1:4){ p5.ac2[i] <- sum(p5.ac1[(i+1):5])  }
for (i in 2:4) {
  p5.aux[i] ~ dbeta(p5.ac1[i],p5.ac2[i])
  pac[3,5,i] <- (1 - sum(pac[3,5,1:(i-1)])) * p5.aux[i]
 }
pac[3,5,5] <- 1 - sum(pac[3,5,1:4])

#-- define distribution for prevalence from accuracy matrix (assumed on reference
method)
prev[5] <- p[3,5]/pac[3,5,5]
prev[4] <- (p[3,4] - prev[5]*pac[3,5,4])/pac[3,4,4]
prev[3] <- (p[3,3] - prev[4]*pac[3,4,3]- prev[5]*pac[3,5,3])/pac[3,3,3]
prev[2] <- (p[3,2] - prev[3]*pac[3,3,2]- prev[4]*pac[3,4,2]-
prev[5]*pac[3,5,2])/pac[3,2,2]
prev[1] <- (p[3,1] - prev[2]*pac[3,2,1]- prev[3]*pac[3,3,1]- prev[4]*pac[3,4,1]-
prev[5]*pac[3,5,1])
      #check <- step(1-sum(prev[1:4]))

### constraint restricting prevalence to be internally consistent with
probabilities
# -- calculate bounds for prevalence (biopsy methods 1 to 3)
for (i in 1:5) { ind5[i] <- equals(rank(p[1:5,5], i),5)   }  # equals(rank(x, i),k)
#1 if ith element of x is kth, 0 otherwise
max.p[5] <- inprod2(ind5[],p[1:5,5])

for (i in 1:5) { aux4[i] <- sum(p[i,4:5]) }
```

```
for (i in 1:5) { ind4[i] <- equals(rank(aux4[1:5], i),5) }
max.p[4] <- inprod2(ind4[],aux4[]) - max.p[5]

for (i in 1:5) { aux3[i] <- sum(p[i,3:5])}
for (i in 1:5) { ind3[i] <- equals(rank(aux3[1:5], i),5) }
max.p[3] <- inprod2(ind3[],aux3[]) - sum(max.p[4:5])

for (i in 1:5) { aux2[i] <- sum(p[i,2:5])}
for (i in 1:5) { ind2[i] <- equals(rank(aux2[1:5], i),5)}
max.p[2] <- inprod2(ind2[],aux2[]) - sum(max.p[3:5])

max.p[1] <- 1- sum(max.p[2:5])

# -- apply constraint on prevalence distribution so that it is consistent with
marginals
for (i in 2:5) {
       z[i] <- 1
       z[i] ~ dbern(constraint[i])
       constraint[i] <- step(prev[i]-max.p[i])*step(1-prev[i])
       }

# -- apply constraint on prevalence distribution to select dist internally coherent
zz <- 1
zz ~ dbern(constraintz)
constraintz <- step(1-sum(prev[2:5]))

####### determining accuracy matrix for COG  #######

## GG2
       pac[1,2,2] <- (p[1,2]-prev[5]*pac[1,5,2]-prev[4]*pac[1,4,2]-
prev[3]*pac[1,3,2])/prev[2]
       pac[1,2,1] <- 1- pac[1,2,2]

## GG3
       pac[1,3,3] <- (p[1,3]-prev[5]*pac[1,5,3]-prev[4]*pac[1,5,4])/prev[3]
                           p31.ac2.aux[1] ~ dbeta(1, 2)
       pac[1,3,1] <- (1-pac[1,3,3])*p31.ac2.aux[1]
       pac[1,3,2] <- 1-pac[1,3,1]-pac[1,3,3]

## GG4
       pac[1,4,4] <- (p[1,4]-prev[5]*pac[1,5,4])/prev[4]
              p41.ac2.aux[1] ~ dbeta(1, 2)
       pac[1,4,1] <- (1-pac[1,4,4])*p41.ac2.aux[1]
              p34.ac2.aux[1] ~ dbeta(1, 1)
       pac[1,4,2] <- (1 - pac[1,4,1]-pac[1,4,4]) * p34.ac2.aux[1]
       pac[1,4,3] <- 1 - sum(pac[1,4,1:2]) - pac[1,4,4]

## GG5
       pac[1,5,5] <- p[1,5]/prev[5]
              p35.ac1.aux[1] ~ dbeta(1, 3)
       pac[1,5,1] <- (1-pac[1,5,5])*p35.ac1.aux[1]
              p35.ac2.aux[1] ~ dbeta(1, 2)
       pac[1,5,2] <- (1-pac[1,5,1]-pac[1,5,5])*p35.ac2.aux[1]
              p35.ac3.aux[1] ~ dbeta(1, 1)
       pac[1,5,3] <- (1-sum(pac[1,5,1:2])-pac[1,5,5])*p35.ac3.aux[1]
       pac[1,5,4] <- 1-sum(pac[1,5,1:3]) - pac[1,5,5]


####### determining accuracy matrix for COG + SB and SOFT +SB, k=4 and k=5 respec
#######
for (k in 4:5) {
## GG2
       pac[k,2,2] <- (p[k,2]-prev[5]*pac[k,5,2]-prev[4]*pac[k,4,2]-
prev[3]*pac[k,3,2])/prev[2]
       pac[k,2,1] <- 1- pac[k,2,2]

## GG3
```

```
        pac[k,3,3] <- (p[k,3]-prev[5]*pac[k,5,3]-prev[4]*pac[k,5,4])/prev[3]
                              p31.ac2.aux[k] ~ dbeta(1, 2)
        pac[k,3,1] <- (1-pac[k,3,3])*p31.ac2.aux[k]
        pac[k,3,2] <- 1-pac[k,3,1]-pac[k,3,3]

## GG4
        pac[k,4,4] <- (p[k,4]-prev[5]*pac[k,5,4])/prev[4]
               p41.ac2.aux[k] ~ dbeta(1, 2)
        pac[k,4,1] <- (1-pac[k,4,4])*p41.ac2.aux[k]
               p34.ac2.aux[k] ~ dbeta(1, 1)
        pac[k,4,2] <- (1 - pac[k,4,1]-pac[k,4,4]) * p34.ac2.aux[k]
        pac[k,4,3] <- 1 - sum(pac[k,4,1:2]) - pac[k,4,4]

## GG5
        pac[k,5,5] <- p[k,5]/prev[5]
               p35.ac1.aux[k] ~ dbeta(1, 3)
        pac[k,5,1] <- (1-pac[k,5,5])*p35.ac1.aux[k]
               p35.ac2.aux[k] ~ dbeta(1, 2)
        pac[k,5,2] <- (1-pac[k,5,1]-pac[k,5,5])*p35.ac2.aux[k]
               p35.ac3.aux[k] ~ dbeta(1, 1)
        pac[k,5,3] <- (1-sum(pac[k,5,1:2])-pac[k,5,5])*p35.ac3.aux[k]
        pac[k,5,4] <- 1-sum(pac[k,5,1:3]) - pac[k,5,5]
        }

## constraints  0<x<1
for ( i in 2:5) {
        for (j in 1:i) {
               z1[1, i,j] <- 1
               z1[1, i,j] ~ dbern(constraintt1[1, i,j])
        constraintt1[1,i,j] <- step(1-pac[1,i,j])*step(pac[1,i,j])
        }
}
for (k in 4:5) {
        for ( i in 2:5) {
               for (j in 1:i) {
                      z1[k, i,j] <- 1
                      z1[k, i,j] ~ dbern(constraintt1[k, i,j])
               constraintt1[k, i,j] <- step(1-pac[k,i,j])*step(pac[k,i,j])
                      }
               }
        }
}
```

Data

ns = number of studies
nt = number of treatments
nc = number of categories
nsX = number of studies of type X
ncX = number of categories in studies type X

T1=cog
T2=SB
T3=fus
T4=cog+SB
T5=fus+SB

list(ns=2, nt=5, nc=5, nsA=4, ncA=2, nsB=5, ncB=3, nsC=2, ncC=4,
#b1=c(0.379032,0.153226,0.209677,0.157258,0.100806), # PAIREDCAP baseline probs - cognitive
#b1=c(0.286290,0.173387,0.282258,0.161290,0.096774), # PAIREDCAP baseline probs - software (Artemis)
b1=c(0.468864,0.164835,0.197802,0.105311,0.063187), # Filson (naive only) baseline probs - software (Artemis)
#b1=c(0.686792,0.086792,0.101887,0.077830,0.046698), # Filson (prior neg) baseline probs - software (Artemis)
b2=c(0.355311,0.219780,0.223443,0.118321,0.083144) # Filson (naive only) baseline probs - software + SB (Artemis) [split by SB proportion in PAIREDCAP]
#b2=c(0.584906,0.150943,0.116981,0.086433,0.060737) # Filson (prior neg) baseline probs - software + SB (Artemis) [split by SB proportion in PAIREDCAP]

#b2=c(0.355311,0.219780,0.223443,0.125916,0.075549) # Filson (naive only) baseline probs - software + SB (Artemis) [split by Artemis proportion in PAIREDCAP]
)

na[]	t[,1]	t[,2]	t[,3]	M[,1]	M[,2]	M[,3]	y[,1,1]	y[,2,1]	y[,3,1]	y[,1,2]	y[,2,2]	
	y[,3,2]	y[,1,3]	y[,2,3]	y[,3,3]	y[,1,4]	y[,2,4]	y[,3,4]	y[,1,5]	y[,2,5]	y[,3,5]	#	study ID
Study type												
3	1	2	3	248	248	248	94	52	71	38	46	43
	52	87	70	39	37	40	25	26	24	#	PAIREDCAP	
all												
2	4	5	NA	100	99	NA	69	55	NA	19	25	NA
	6	13	NA	5	3	NA	1	3	NA	#	Izadpanahi	
all												
3	2	3	5	169	169	169	53	49	36	116	120	133
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Wajswol 2020	
A												
3	1	2	4	75	75	75	41	35	32	34	40	43
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Thangarasu	
2021	A											
3	1	2	4	63	63	63	30	33	25	33	30	38
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Kulis 2020	
A												
2	1	3	NA	88	88	NA	57	48	NA	31	40	NA
	NA	NA	NA	NA	NA	NA	NA	NA	NA	#	Cornud	A
2	1	3	NA	78	79	NA	44	40	NA	8	12	NA
	26	27	NA	NA	NA	NA	NA	NA	NA	#	FUTURE	B
2	1	3	NA	125	125	NA	85	80	NA	16	16	NA
	24	29	NA	NA	NA	NA	NA	NA	NA	#	PROFUS	B
3	2	3	5	74	74	74	41	39	32	12	10	13
	21	25	29	NA	NA	NA	NA	NA	NA	#	Albisinni 2018	
B												
3	2	3	5	191	191	191	103	106	85	36	25	34
	52	60	72	NA	NA	NA	NA	NA	NA	#	Fourcade 2018	
B												
3	1	2	4	111	111	111	69	81	65	19	9	20
	23	21	26	NA	NA	NA	NA	NA	NA	#	Gomez-Ortiz	
2022	B											
3	2	3	5	48	48	48	23	20	16	11	11	13
	10	13	13	4	4	6	NA	NA	NA	#	Alberts 2018	
(all men)	C											
3	2	3	5	538	538	538	294	310	252	114	68	100
	74	81	92	56	79	94	NA	NA	NA	#	Filson 2016	
C												

END

#extension
list(

#B.y = c(94,38,52,39,25), B.M=248, #PAIREDCAP baseline cognitive
#B.y = c(71, 43, 70, 40, 24), B.M =248, #PAIREDCAP baseline fusion
B.y = c(128, 45, 54, 29, 17), B.M =273, ## FUS -- filson naive
B.z = c(97, 60, 61, 32, 23), B.Mz=273, ## FUS+SB -- filson naive
#B.y = c(182, 23, 27, 21, 12), B.M =265 ## FUS -- filson prior negative
#B.z = c(155, 40, 31, 23, 16), B.Mz=265, ## FUS+SB -- filson prior negative
# y2.ac = c(11,14) ,
# y3.ac = c(13,1,15),
# y4.ac= c(1,1,2,16),
# y5.ac= c(0,0,1,1,25),
# M.ac=c(25,29,20,27) # conditional accuracy matrix Zhou
y2.ac = c(24,11) ,
y3.ac = c(21,17,45),
y4.ac= c(10,2,9,24),
y5.ac= c(6,2,7,8,36),
M.ac=c(35,83,45,59) # conditional accuracy matrix Mortezavi
)

Initial values

```
#chain 1
list( a = structure(.Data = c(  NA,0,0,0,0,          NA,0,0,0,0,          NA,0,NA,NA,NA,          NA,0,NA,NA,NA,
NA,0,NA,NA,NA,          NA,0,NA,NA,NA,          NA,0,0,NA,NA,          NA,0,0,NA,NA,
NA,0,0,NA,NA,          NA,0,0,NA,NA,          NA,0,0,0,NA,          NA,0,0,0,NA),
.Dim = c(13,5)),
d = structure(.Data = c( NA,NA,NA,NA,NA,          NA,0,0,0,0,          NA,0,0,0,0,          NA,0,0,0,0,
NA,0,0,0,0),
.Dim = c(5,5)),
dA = structure(.Data = c(NA,NA,          NA,0,          NA,0,          NA,0,          NA,0),
.Dim = c(5,2)),
dB = structure(.Data = c(NA,NA,NA,          NA,NA,0,          NA,NA,0,          NA,NA,0,          NA,NA,0),
.Dim = c(5,3)),
dC = structure(.Data = c(NA,NA,NA,NA,          NA,NA,NA,0,          NA,NA,NA,0,          NA,NA,NA,0,
NA,NA,NA,0),
.Dim = c(5,4))
)

#chain 2
list( a = structure(.Data = c(  NA,2,-.5,1,-1,          NA,2,3,1,2,          NA,-2,NA,NA,NA,          NA,-2,NA,NA,NA,
NA,-2,NA,NA,NA,          NA,-2,NA,NA,NA,          NA,-2,1,NA,NA,          NA,1,-2,NA,NA,          NA,-2,1,NA,NA,
NA,-2,1,NA,NA,          NA,-2,1,NA,NA,          NA,.7,-2,-1,NA,          NA,1,-2,2,NA),
.Dim = c(13,5)),
d = structure(.Data = c( NA,NA,NA,NA,NA,          NA,-1,-2,1,2,          NA,.5,-2,-1,1,          NA,2,-2,.5,-2,
NA,1,2,1,-2),
.Dim = c(5,5)),
dA = structure(.Data = c(NA,NA,          NA,-2,          NA,2,          NA,1,          NA,2),
.Dim = c(5,2)),
dB = structure(.Data = c(NA,NA,NA,          NA,NA,1,          NA,NA,-1,          NA,NA,-2,          NA,NA,-1),
.Dim = c(5,3)),
dC = structure(.Data = c(NA,NA,NA,NA,          NA,NA,NA,2,          NA,NA,NA,.7,          NA,NA,NA,-.5,
NA,NA,NA,-2),
.Dim = c(5,4))
)
```

Additional results

Influence of the use of constraints on the network meta-analysis estimates

Comparison of inferences on distribution of test results with the synthesis code used in isolation and the synthesis code including the extension.

Table 84: Influence of the model extension on inferences over the probability of test results for biopsy naïve. Main analysis.

Model	Synthesis model	Extended synthesis and accuracy model		
Assumptions over baseline probability	Deterministic	Deterministic	Probabilistic	Probabilistic
Calculation of distribution of test results	Directly from synthesis	Directly from synthesis	Directly from synthesis	Back calculated from prevalence and accuracy matrix
Biopsy method **Category**				
Network 1				
Cognitive fusion **No cancer**	0.552 [0.475,0.624]	0.545 [0.477,0.612]	0.531 [0.446,0.616]	0.513 [0.414,0.610]
Cognitive fusion **1**	0.174 [0.132,0.223]	0.177 [0.135,0.225]	0.185 [0.129,0.249]	0.185 [0.132,0.245]
Cognitive fusion **2**	0.118 [0.081,0.164]	0.121 [0.085,0.165]	0.120 [0.078,0.173]	0.139 [0.070,0.217]
Cognitive fusion **3**	0.094 [0.058,0.143]	0.095 [0.059,0.142]	0.098 [0.055,0.160]	0.097 [0.056,0.159]
Cognitive fusion **4 or 5**	0.062 [0.034,0.104]	0.062 [0.035,0.099]	0.066 [0.032,0.117]	0.065 [0.032,0.115]
Software fusion **No cancer**	0.469	0.469	0.457 [0.403,0.509]	0.457 [0.403,0.513]
Software fusion **1**	0.165	0.165	0.172 [0.137,0.212]	0.173 [0.137,0.214]
Software fusion **2**	0.198	0.198	0.195 [0.159,0.236]	0.196 [0.157,0.233]
Software fusion **3**	0.105	0.105	0.109 [0.080,0.146]	0.108 [0.079,0.144]
Software fusion **4 or 5**	0.063	0.063	0.067 [0.045,0.095]	0.066 [0.043,0.095]
Network 2				
Combined cognitive fusion and systematic biopsy **No cancer**	0.402 [0.210,0.559]	0.451 [0.345,0.562]	0.455 [0.343,0.570]	0.460 [0.335,0.583]
Combined cognitive fusion and systematic biopsy **1**	0.211 [0.102,0.326]	0.246 [0.16,0.349]	0.249 [0.156,0.362]	0.250 [0.152,0.356]
Combined cognitive fusion and systematic biopsy **2**	0.109 [0.031,0.238]	0.136 [0.054,0.256]	0.134 [0.052,0.255]	0.127 [0.034,0.261]
Combined cognitive fusion and systematic biopsy **3**	0.241 [0.058,0.586]	0.135 [0.048,0.245]	0.130 [0.045,0.230]	0.131 [0.046,0.231]
Combined cognitive fusion and systematic biopsy **4 or 5**	0.037 [0.001,0.172]	0.033 [0.001,0.107]	0.032 [0.001,0.107]	0.033 [0.001,0.107]
Combined software fusion and systematic biopsy **No cancer**	0.355	0.355	0.359 [0.305,0.413]	0.346 [0.274,0.408]
Combined software fusion and systematic biopsy **1**	0.220	0.220	0.223 [0.178,0.273]	0.222 [0.177,0.273]
Combined software fusion and systematic biopsy **2**	0.223	0.223	0.221 [0.177,0.270]	0.234 [0.170,0.313]
Combined software fusion and systematic biopsy **3**	0.118	0.118	0.115 [0.082,0.154]	0.116 [0.084,0.153]
Combined software fusion and systematic biopsy **4 or 5**	0.083	0.083	0.082 [0.055,0.114]	0.082 [0.054,0.114]

Results from this comparison show that for network 1 the structural extension model does not significantly influence synthesis estimates. For network 2, estimates of category 4 for the non-

reference treatment (combined cognitive fusion and systematic biopsy) are reduced in the extended model, which suggests a conflict between the structural extension (including data sources added) and the uncertainty derived from the multinomial log odds model implemented in the synthesis. For this category, there is only one study providing a direct comparison of combined software vs. combined cognitive fusion with very few patients classified in categories 4 or 5,[82] providing very sparse information. This study reports a proportion of 5% of test results in category 4 with combined cognitive, vs. 3% in combined software fusion. Therefore, uncertainty is very wide for this category and the constrained model restricts the distribution of this category the most.

Table 85: Influence of the model extension on inferences over the probability of test results for previous negative biopsy. Subgroup analysis.

Model		Synthesis model	Extended synthesis and accuracy model		
Assumptions over baseline probability		Deterministic	Deterministic	Probabilistic	Probabilistic
Calculation of distribution of test results		Directly from synthesis	Directly from synthesis	Directly from synthesis	Back calculated from prevalence and accuracy matrix
Biopsy method	**ISUP grade**				
Network 1					
Cognitive fusion	**No cancer**	0.750 [0.688,0.803]	0.744 [0.686,0.798]	0.719 [0.643,0.788]	0.703 [0.618,0.776]
	1	0.085 [0.062,0.114]	0.088 [0.065,0.12]	0.105 [0.069,0.153]	0.105 [0.071,0.155]
	2	0.057 [0.038,0.082]	0.058 [0.04,0.079]	0.061 [0.039,0.091]	0.077 [0.035,0.138]
	3	0.065 [0.039,0.101]	0.066 [0.038,0.101]	0.068 [0.036,0.116]	0.068 [0.033,0.120]
	4 or 5	0.043 [0.023,0.074]	0.044 [0.024,0.07]	0.047 [0.021,0.086]	0.046 [0.021,0.085]
Software fusion	**No cancer**	0.687	0.687	0.661 [0.611,0.710]	0.661 [0.611,0.709]
	1	0.087	0.087	0.103 [0.075,0.137]	0.103 [0.076,0.135]
	2	0.102	0.102	0.107 [0.082,0.136]	0.107 [0.082,0.136]
	3	0.078	0.078	0.080 [0.055,0.110]	0.080 [0.055,0.111]
	4 or 5	0.047	0.047	0.049 [0.031,0.073]	0.049 [0.031,0.073]
Network 2					
Combined cognitive fusion and systematic biopsy	**No cancer**	0.615 [0.382,0.76]	0.658 [0.557,0.75]	0.664 [0.56,0.761]	0.659 [0.561,0.752]
	1	0.135 [0.073,0.208]	0.152 [0.098,0.221]	0.155 [0.093,0.240]	0.157 [0.096,0.241]
	2	0.053 [0.015,0.120]	0.067 [0.025,0.135]	0.065 [0.023,0.136]	0.067 [0.015,0.152]
	3	0.171 [0.036,0.468]	0.095 [0.03,0.181]	0.090 [0.028,0.163]	0.091 [0.027,0.165]
	4 or 5	0.027 [0.000,0.125]	0.029 [0.001,0.099]	0.027 [0.001,0.087]	0.027 [0.001,0.081]
Combined software fusion and	**No cancer**	0.585	0.585	0.591 [0.536,0.647]	0.583 [0.513,0.649]
	1	0.151	0.151	0.154 [0.114,0.201]	0.155 [0.114,0.198]

2	0.117	0.117	0.113 [0.080,0.147]	0.120 [0.074,0.181]
3	0.086	0.086	0.084 [0.056,0.118]	0.083 [0.057,0.117]
4 or 5	0.061	0.061	0.058 [0.036,0.084]	0.058 [0.037,0.084]

Sensitivity analysis

Table 86: Influence of the model extension on inferences over the probability of test results for biopsy naïve. Sensitivity analysis to baseline distribution (PAIREDCAP's baseline, Mortezavi's accuracy).

	Model	Synthesis model	Extended synthesis and accuracy model		
	Assumptions over baseline probability	Deterministic	Deterministic	Probabilistic	Probabilistic
	Calculation of distribution of test results	Directly from synthesis	Directly from synthesis	Directly from synthesis	Back calculated from prevalence and accuracy matrix
	ISUP				
Network 1					
Cognitive fusion	**No cancer**	0.363 [0.294,0.435]	0.364 [0.293,0.44]	0.392 [0.308,0.482]	0.368 [0.248,0.473]
	1	0.197 [0.148,0.255]	0.199 [0.152,0.253]	0.192 [0.133,0.263]	0.191 [0.140,0.256]
	2	0.182 [0.130,0.245]	0.184 [0.130,0.25]	0.169 [0.111,0.242]	0.196 [0.101,0.306]
	3	0.156 [0.100,0.226]	0.152 [0.095,0.219]	0.147 [0.084,0.232]	0.145 [0.079,0.228]
	4 or 5	0.102 [0.057,0.167]	0.102 [0.059,0.154]	0.100 [0.051,0.176]	0.101 [0.052,0.176]
Software fusion	**No cancer**	0.286	0.286	0.313 [0.271,0.360]	0.314 [0.271,0.362]
	1	0.173	0.173	0.170 [0.135,0.209]	0.169 [0.137,0.207]
	2	0.282	0.282	0.262 [0.218,0.310]	0.263 [0.218,0.308]
	3	0.161	0.161	0.158 [0.117,0.203]	0.157 [0.117,0.204]
	4 or 5	0.097	0.097	0.097 [0.064,0.137]	0.098 [0.064,0.140]

Table 87: Influence of the model extension on inferences over the probability of test results for biopsy naïve. Sensitivity analysis to accuracy matrix (Filson's baseline, Zhou's accuracy).

Model	Synthesis model	Extended synthesis and accuracy model		
Assumptions over baseline probability	Deterministic	Deterministic	Probabilistic	Probabilistic
Calculation of distribution of test results	Directly from synthesis	Directly from synthesis	Directly from synthesis	Back calculated from prevalence

					and accuracy matrix
	ISUP				
Network 1					
Cognitive fusion	**No cancer**	0.552 [0.475,0.624]	0.555 [0.488,0.626]	0.531 [0.445,0.621]	0.525 [0.433,0.620]
	1	0.174 [0.132,0.223]	0.179 [0.132,0.232]	0.191 [0.132,0.264]	0.190 [0.131,0.256]
	2	0.118 [0.081,0.164]	0.117 [0.082,0.161]	0.112 [0.070,0.169]	0.122 [0.062,0.201]
	3	0.094 [0.058,0.143]	0.093 [0.059,0.137]	0.099 [0.055,0.160]	0.098 [0.053,0.158]
	4 or 5	0.062 [0.034,0.104]	0.056 [0.032,0.083]	0.066 [0.034,0.109]	0.065 [0.033,0.106]
Software fusion	**No cancer**	0.469	0.469	0.449 [0.396,0.503]	0.450 [0.400,0.509]
	1	0.165	0.165	0.176 [0.139,0.220]	0.175 [0.140,0.217]
	2	0.198	0.198	0.189 [0.149,0.234]	0.189 [0.147,0.236]
	3	0.105	0.105	0.112 [0.082,0.149]	0.112 [0.082,0.146]
	4 or 5	0.063	0.063	0.075 [0.053,0.105]	0.075 [0.053,0.103]

Detailed results for subgroup analysis on previous negative biopsy individuals

Table 88: Distribution of test results, conditional accuracy and prevalence probabilities (mean and 95% CrI) according to ISUP Grade (D) for individuals with a previous negative biopsy. Diagnostic accuracy extension to the evidence synthesis model. Results of subgroup analysis.

Network 1

		0.703 [0.618,0.776]	0.105 [0.071,0.155]	0.077 [0.035,0.138]	0.068 [0.033,0.12]	0.046 [0.021,0.085]		0.661 [0.611,0.709]	0.103 [0.076,0.135]	0.107 [0.082,0.136]	0.080 [0.055,0.111]	0.049 [0.031,0.073]
		cognitive fusion						**software fusion**				
	ISUP	**No cancer**	**1**	**2**	**3**	**4 or 5**		**No cancer**	**1**	**2**	**3**	**4 or 5**
0.428 [0.259,0.529]	**No cancer**	1	0	0	0	0		1	0	0	0	0
0.224 [0.138,0.39]	**1**	0.857 [0.575,0.995]	0.143 [0.005,0.425]	0	0	0		0.698 [0.557,0.826]	0.302 [0.174,0.443]	0	0	0
0.132 [0.091,0.188]	**2**	0.324 [0.026,0.722]	0.374 [0.066,0.718]	0.302 [0.049,0.557]	0	0		0.264 [0.183,0.35]	0.201 [0.13,0.288]	0.536 [0.436,0.626]	0	0
0.131 [0.079,0.199]	**3**	0.195 [0.005,0.571]	0.130 [0.003,0.428]	0.208 [0.007,0.579]	0.466 [0.181,0.793]	0		0.225 [0.129,0.344]	0.054 [0.011,0.125]	0.195 [0.103,0.301]	0.526 [0.404,0.648]	0
0.085 [0.053,0.127]	**4 or 5**	0.136 [0.002,0.439]	0.109 [0.002,0.324]	0.115 [0.003,0.363]	0.093 [0.004,0.315]	0.547 [0.283,0.902]		0.110 [0.044,0.193]	0.046 [0.01,0.114]	0.122 [0.055,0.216]	0.141 [0.071,0.226]	0.581 [0.462,0.694]

Network 2

		0.659 [0.561,0.752]	0.157 [0.096,0.241]	0.067 [0.015,0.152]	0.091 [0.027,0.165]	0.027 [0.001,0.081]		0.583 [0.513,0.649]	0.155 [0.114,0.198]	0.120 [0.074,0.181]	0.083 [0.057,0.117]	0.058 [0.037,0.084]
		Combined cognitive fusion and systematic biopsy						**Combined software fusion and systematic biopsy**				
	ISUP	**No cancer**	**1**	**2**	**3**	**4 or 5**		**1**	**2**	**3**	**4**	**5**
0.428 [0.259,0.529]	**No cancer**	1	0	0	0	0		1	0	0	0	0
0.224 [0.138,0.39]	**1**	0.693 [0.251,0.984]	0.307 [0.016,0.749]	0	0	0		0.497 [0.153,0.845]	0.503 [0.155,0.847]	0	0	0
0.132 [0.091,0.188]	**2**	0.272 [0.009,0.709]	0.448 [0.058,0.864]	0.281 [0.017,0.791]	0	0		0.087 [0.002,0.315]	0.176 [0.012,0.46]	0.736 [0.408,0.971]	0	0
0.131 [0.079,0.199]	**3**	0.136 [0.001,0.515]	0.121 [0.002,0.474]	0.140 [0.001,0.515]	0.603 [0.126,0.982]	0		0.138 [0.003,0.439]	0.132 [0.002,0.44]	0.131 [0.003,0.418]	0.600 [0.298,0.938]	0
0.085 [0.053,0.127]	**4 or 5**	0.207 [0.003,0.629]	0.196 [0.004,0.577]	0.147 [0.004,0.546]	0.135 [0.003,0.434]	0.315 [0.014,0.899]		0.071 [0.001,0.251]	0.074 [0.001,0.284]	0.073 [0.001,0.266]	0.083 [0.001,0.311]	0.699 [0.367,0.966]

Table 89 Distribution of test results, joint probability matrix and prevalence probabilities (mean and 95% CrI) according to ISUP Grade (D) for individuals with a previous negative biopsy. Diagnostic accuracy extension to the evidence synthesis model. Results of subgroup analysis.

Network 1

		(distribution of test results)					(distribution of test results)				
		0.703 [0.618,0.776]	0.105 [0.071,0.155]	0.077 [0.035,0.138]	0.068 [0.033,0.12]	0.046 [0.021,0.085]	0.661 [0.611,0.709]	0.103 [0.076,0.135]	0.107 [0.082,0.136]	0.080 [0.055,0.111]	0.049 [0.031,0.073]
		cognitive fusion					**software fusion**				
		(joint probability matrix)					(joint probability matrix)				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.428 [0.259,0.529]	**No cancer**	0.428 [0.259,0.529]	0	0	0	0	0.428 [0.259,0.529]	0	0	0	0
0.224 [0.138,0.39]	**1**	0.194 [0.092,0.363]	0.03 [0.001,0.085]	0	0	0	0.159 [0.082,0.313]	0.065 [0.037,0.098]	0	0	0
0.132 [0.091,0.188]	**2**	0.044 [0.003,0.103]	0.049 [0.009,0.095]	0.04 [0.006,0.076]	0	0	0.035 [0.02,0.057]	0.026 [0.014,0.043]	0.071 [0.047,0.104]	0	0
0.131 [0.079,0.199]	**3**	0.026 [0.001,0.087]	0.017 [0,0.056]	0.028 [0.001,0.085]	0.06 [0.02,0.113]	0	0.03 [0.013,0.061]	0.007 [0.001,0.019]	0.026 [0.011,0.047]	0.068 [0.041,0.101]	0
0.085 [0.053,0.127]	**4 or 5**	0.012 [0,0.041]	0.009 [0,0.029]	0.01 [0,0.031]	0.008 [0,0.03]	0.046 [0.021,0.085]	0.009 [0.003,0.018]	0.004 [0.001,0.011]	0.01 [0.004,0.021]	0.012 [0.005,0.023]	0.049 [0.031,0.073]

Network 2

		(distribution of test results)					(distribution of test results)				
		0.659 [0.561,0.752]	0.157 [0.096,0.241]	0.067 [0.015,0.152]	0.091 [0.027,0.165]	0.027 [0.001,0.081]	0.583 [0.513,0.649]	0.155 [0.114,0.198]	0.120 [0.074,0.181]	0.083 [0.057,0.117]	0.058 [0.037,0.084]
		Combined cognitive fusion and systematic biopsy					**Combined software fusion and systematic biopsy**				
		(joint probability matrix)					(joint probability matrix)				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.121 [0.007,0.238]	**No cancer**	0.428 [0.259,0.529]	0	0	0	0	0.428 [0.259,0.529]	0	0	0	0
0.318 [0.212,0.452]	**1**	0.159 [0.041,0.344]	0.065 [0.004,0.145]	0	0	0	0.118 [0.025,0.3]	0.106 [0.039,0.16]	0	0	0
0.262 [0.193,0.341]	**2**	0.037 [0.001,0.101]	0.059 [0.007,0.122]	0.036 [0.002,0.11]	0	0	0.012 [0,0.048]	0.024 [0.001,0.072]	0.096 [0.054,0.136]	0	0
0.183 [0.119,0.265]	**3**	0.018 [0,0.072]	0.016 [0,0.063]	0.018 [0,0.071]	0.079 [0.015,0.154]	0	0.019 [0,0.074]	0.018 [0,0.064]	0.018 [0,0.068]	0.075 [0.04,0.115]	0

0.116 [0.077,0.174]	**4 or 5**	0.018 [0,0.057]	0.017 [0,0.056]	0.012 [0,0.045]	0.012 [0,0.044]	0.027 [0.001,0.081]		0.007 [0,0.027]	0.007 [0,0.029]	0.007 [0,0.027]	0.008 [0,0.031]	0.058 [0.037,0.084]

Detailed results of sensitivity analyses

Table 90: Distribution of test results, conditional accuracy and prevalence probabilities (mean and 95% CrI) according to ISUP Grade for biopsy naïve individuals. Diagnostic accuracy extension to the evidence synthesis model. Results of sensitivity analysis using PAIREDCAP baseline and Mortezavi accuracy.

Network 1

		0.368 [0.248,0.473]	0.191 [0.14,0.256]	0.196 [0.101,0.306]	0.145 [0.079,0.228]	0.101 [0.052,0.176]		0.314 [0.271,0.362]	0.169 [0.137,0.207]	0.263 [0.218,0.308]	0.157 [0.117,0.204]	0.098 [0.064,0.14]
		cognitive fusion						**software fusion**				
	ISUP	**No cancer**	**1**	**2**	**3**	**4 or 5**		**No cancer**	**1**	**2**	**3**	**4 or 5**
0.031 [0.001,0.092]	**No cancer**	1	0	0	0	0		1	0	0	0	0
0.226 [0.163,0.319]	**1**	0.754 [0.355,0.990]	0.246 [0.010,0.645]	0	0	0		0.634 [0.506,0.773]	0.366 [0.227,0.494]	0	0	0
0.322 [0.222,0.42]	**2**	0.300 [0.025,0.642]	0.288 [0.071,0.532]	0.412 [0.133,0.687]	0	0		0.229 [0.155,0.305]	0.197 [0.123,0.276]	0.575 [0.484,0.67]	0	0
0.252 [0.154,0.37]	**3**	0.175 [0.007,0.520]	0.121 [0.004,0.402]	0.183 [0.006,0.492]	0.521 [0.27,0.862]	0		0.191 [0.105,0.294]	0.058 [0.013,0.121]	0.222 [0.124,0.331]	0.530 [0.412,0.659]	0
0.169 [0.104,0.254]	**4 or 5**	0.121 [0.004,0.392]	0.086 [0.001,0.285]	0.105 [0.003,0.342]	0.088 [0.002,0.31]	0.599 [0.324,0.920]		0.102 [0.041,0.184]	0.044 [0.009,0.107]	0.127 [0.06,0.216]	0.144 [0.069,0.236]	0.583 [0.449,0.700]

Table 91: Distribution of test results, joint probability matrix and prevalence probabilities (mean and 95% CrI) according to ISUP Grade for biopsy naïve individuals. Diagnostic accuracy extension to the evidence synthesis model. Results of sensitivity analysis using PAIREDCAP baseline and Mortezavi accuracy.

Network 1		(distribution of test results)					(distribution of test results)				
		0.368 [0.248,0.473]	0.191 [0.14,0.256]	0.196 [0.101,0.306]	0.145 [0.079,0.228]	0.101 [0.052,0.176]	0.314 [0.271,0.362]	0.169 [0.137,0.207]	0.263 [0.218,0.308]	0.157 [0.117,0.204]	0.098 [0.064,0.14]
		cognitive fusion					**software fusion**				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.031 [0.001,0.092]	**No cancer**	0.031 [0.001,0.092]	0	0	0	0	0.031 [0.001,0.092]	0	0	0	0
0.226 [0.163,0.319]	**1**	0.171 [0.07,0.271]	0.055 [0.002,0.151]	0	0	0	0.143 [0.098,0.202]	0.083 [0.043,0.129]	0	0	0
0.322 [0.222,0.42]	**2**	0.099 [0.006,0.235]	0.092 [0.021,0.173]	0.131 [0.042,0.21]	0	0	0.074 [0.042,0.114]	0.063 [0.036,0.097]	0.185 [0.125,0.246]	0	0
0.252 [0.154,0.37]	**3**	0.045 [0.001,0.144]	0.03 [0.001,0.095]	0.047 [0.001,0.142]	0.13 [0.055,0.219]	0	0.048 [0.021,0.088]	0.015 [0.003,0.035]	0.057 [0.024,0.1]	0.132 [0.081,0.181]	0
0.169 [0.104,0.254]	**4 or 5**	0.021 [0.001,0.079]	0.015 [0,0.054]	0.018 [0,0.064]	0.015 [0,0.056]	0.101 [0.052,0.176]	0.018 [0.006,0.038]	0.008 [0.001,0.02]	0.022 [0.008,0.048]	0.025 [0.009,0.049]	0.098 [0.064,0.14]

Table 92: Distribution of test results, conditional accuracy and prevalence probabilities (mean and 95% CrI) according to ISUP Grade for biopsy naïve individuals. Diagnostic accuracy extension to the evidence synthesis model. Results of sensitivity analysis using using Filson's baseline and Zhou's accuracy.

Network 1

		0.525 [0.433,0.62]	0.190 [0.131,0.256]	0.122 [0.062,0.201]	0.098 [0.053,0.158]	0.065 [0.033,0.106]	0.450 [0.400,0.509]	0.175 [0.140,0.217]	0.189 [0.147,0.236]	0.112 [0.082,0.146]	0.075 [0.053,0.103]
		cognitive fusion					**software fusion**				
	ISUP	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.170 [0.023,0.280]	**No cancer**	1	0	0	0	0	1	0	0	0	0
0.279 [0.196,0.400]	**1**	0.787 [0.435,0.992]	0.213 [0.008,0.565]	0	0	0	0.472 [0.321,0.631]	0.528 [0.369,0.679]	0	0	0
0.300 [0.211,0.436]	**2**	0.327 [0.03,0.693]	0.362 [0.065,0.646]	0.312 [0.127,0.518]	0	0	0.415 [0.268,0.569]	0.048 [0.006,0.129]	0.537 [0.392,0.691]	0	0
0.155 [0.109,0.223]	**3**	0.152 [0.003,0.448]	0.113 [0.003,0.367]	0.147 [0.003,0.477]	0.588 [0.292,0.919]	0	0.094 [0.013,0.257]	0.077 [0.013,0.203]	0.144 [0.035,0.31]	0.685 [0.512,0.846]	0
0.095 [0.067,0.136]	**4 or 5**	0.083 [0.001,0.326]	0.074 [0.001,0.278]	0.080 [0.002,0.284]	0.080 [0.001,0.296]	0.683 [0.391,0.975]	0.037 [0.001,0.128]	0.034 [0.001,0.113]	0.075 [0.011,0.189]	0.068 [0.01,0.175]	0.787 [0.65,0.915]

Table 93: Distribution of test results, joint probability matrix and prevalence probabilities (mean and 95% CrI) according to ISUP Grade for biopsy naïve individuals. Diagnostic accuracy extension to the evidence synthesis model. Results of sensitivity analysis using using Filson's baseline and Zhou's accuracy.

Network 1		(distribution of test results)					(distribution of test results)				
		0.525 [0.433,0.62]	0.190 [0.131,0.256]	0.122 [0.062,0.201]	0.098 [0.053,0.158]	0.065 [0.033,0.106]	0.450 [0.400,0.509]	0.175 [0.140,0.217]	0.189 [0.147,0.236]	0.112 [0.082,0.146]	0.075 [0.053,0.103]
		cognitive fusion					**software fusion**				
(prevalence)	**ISUP**	**No cancer**	**1**	**2**	**3**	**4 or 5**	**No cancer**	**1**	**2**	**3**	**4 or 5**
0.170 [0.023,0.280]	**No cancer**	0.170 [0.023,0.28]	0	0	0	0	0.170 [0.023,0.28]	0	0	0	0
0.279 [0.196,0.400]	**1**	0.221 [0.101,0.355]	0.059 [0.002,0.158]	0	0	0	0.134 [0.068,0.24]	0.145 [0.103,0.192]	0	0	0
0.300 [0.211,0.436]	**2**	0.102 [0.009,0.261]	0.106 [0.019,0.194]	0.092 [0.039,0.154]	0	0	0.127 [0.064,0.235]	0.015 [0.001,0.042]	0.158 [0.111,0.205]	0	0
0.155 [0.109,0.223]	**3**	0.024 [0,0.077]	0.018 [0,0.058]	0.023 [0,0.077]	0.09 [0.042,0.153]	0	0.015 [0.002,0.042]	0.012 [0.002,0.033]	0.023 [0.004,0.057]	0.105 [0.074,0.142]	0
0.095 [0.067,0.136]	**4 or 5**	0.008 [0,0.035]	0.007 [0,0.029]	0.008 [0,0.029]	0.008 [0,0.028]	0.065 [0.033,0.106]	0.004 [0,0.013]	0.003 [0,0.012]	0.007 [0.001,0.021]	0.007 [0.001,0.018]	0.075 [0.053,0.103]

APPENDIX 11. MODEL PARAMETERISATION

Treatments distribution

Table 94 Metastatic disease treatment allocation by diagnosed category

Metastatic hormone-sensitive treatment	**Southampton DAR**		**Current DAR**	
ADT alone	50%		50%	
Docetaxel + ADT	36%		9%	
Enzalutamide + ADT	7%		34%	
Apalutamide + ADT	7%		7%	
Metastatic hormone-relapsed treatment	**Previously treated with**			
	ADT alone	**Docetaxel + ADT**	**Enzalutamide + ADT**	**Apalutamide + ADT**
Abiraterone	35%	30%	0%	0%
Docetaxel	10%	25%	60%	60%
Enzalutamide	35%	35%	0%	0%
BSC	20%	10%	40%	40%

ADT, androgen deprivation therapy; BSC, best supportive treatment

Adverse events

Table 95 Biopsy adverse event rates applied in the model

	Adverse event rates			
Adverse event	**LATRUS**	**Source**	**LATP**	**Source**
Mild adverse event	1.31%	Rosario et al., 2012	9.13%	Pepe & Aragona et al., 2013[185] – emergency visits all patients
Non-elective admission*	3.74%	Tamhankar et al., 2020	3.54%	Tamhankar et al., 2020[186]
Death*	0.07%		0.05%	

*Within 28 days of biopsy; LATP, local anaesthesia transperineal; LATRUS, local anaesthesia transrectal ultrasound

HRQoL

Table 96 Parameterisation of biopsy procedural disutility

Adverse event	**Disutility weight**	**Duration of adverse event (days)**	**QALY loss**
Mild adverse event	0.29	3	0.002
Non-elective admission*	0.49	30	0.041
Death*	0.49	30	0.041

*Within 28 days of biopsy; QALY, quality-adjusted life year

Resource use and costs

Patient throughput

We considered the estimates of throughput applied in the Southampton DAR,[126] which assumed 18 weekly and 1,000 annual biopsies (not distinguishing throughput between systematic and targeted biopsies). Clinical advisers to the EAG considered that the annual estimate is likely to overestimate the average total number of biopsies per NHS trust and may be more reflective of a very high throughput centre.

We also examined prostate biopsy activity numbers across all healthcare resource groups (HRGs) in the main schedule of NHS reference costs across three financial years (2018/2019, 2019/2020, 2020/21)[165, 166, 187]for the prostate biopsy currency codes across all HRG data (LB76Z [Transrectal Ultrasound Guided Biopsy of Prostate] and LB77Z [Transperineal Template Biopsy of Prostate]) and contrasted these figures against those reported for the latest available National Prostate Cancer Audit (NPCA) annual report,[4] as illustrated in Table 97. We did not consider earlier versions of the NPCA annual reports due to changes in the reporting style and high level of missing data, which hinder establishing meaningful comparisons across time. We note as a limitation of the NHS reference data that the transperineal biopsy currency code suggests these were transperineal template biopsies, so it is unclear how other types of transperineal biopsies were captured in the dataset.

Table 97 Evidence considered to estimate the patient throughput

	Data source					
	NHS reference costs; all HRG data			NPCA annual report[4]]		
Data collection period	2018/19 financial year[187]	2019/20 financial year[165]	2020/21 financial year[166]	April 2019-March 2020		
Country	England			England	Wales	England & Wales
Biopsy route						
. TP biopsy	39,211	30,451	11,492	20,623	969	21,592
. TR biopsy	2,1424	21,674	22,332	13,756	300	13,756
Total biopsies per year	60,635	52,125	33,824	34,379	1,269	35,348
Estimated annual number of biopsies preceded by an MRI/NHS trust*	52,752	45,349	29,427	29910	1104	30753
Estimated annual number of biopsies/NHS trust*	415	357	232	236	221	235

Estimated annual number of targeted biopsies/NHS trust*	300	258	168	170	160	170

*Or University Health Board if in Wales; HRG, Healthcare Resource Group; NPCA, National Prostate Cancer Audit; TP, transperineal; TRUS, transrectal

Although the NPCA reports data for both England and Wales, the total number of biopsies reported is lower than that reported for a similar period in the main schedule NHS reference costs; this is due to missing data issues. To estimate the average number of biopsies per NHS England trust and/or Welsh University Health board, we assumed the number of institutions from which the NCPA collected data in 2019/20 (127 NHS trusts and 5 University Health Boards).[188] Although clinical guidance has recommended performing a mpMRI before any biopsy is offered at least since 2019,NICE has identified data suggesting that in 2019[189] only 87% of biopsies were preceded by an mpMRI in England and Wales. Thus, we used the 87% estimate (varied in scenario analysis to 100%, Section 2.2.2.1, to explore the impact of complete compliance with clinical guidance) to adjust the average annual number of biopsies by NHS trust. Finally, we estimated the average annual number of targeted biopsies by assuming that 72.6% of biopsies preceded by an mpMRI had a Likert or PI-RADS score of at least 3, as this is the threshold at which targeted biopsy is recommended. The 72.6% was obtained by pooling the proportion of patients in two relevant RCTs (71.8% in PROMIS [UK] and 72.6% PRECISION [11 countries])[19, 136] who had a mpMRI result of at least 3 (Likert or PI-RADS).

The evidence considered suggests the average annual number of targeted biopsies (alone or in combination with systematic biopsy) per NHS trust in England is in a range within 168 and 300. However, the two latest data cuts of NHS reference costs[165, 166, 187] are likely to be affected to some extend the impact of COVID related constraints on NHS service provision. Therefore, we consider that the expected patient throughput is likely to be closer to the upper bound of the estimated range and consider an annual throughput of 300 targeted biopsies in the base-case analysis.

Biopsy procedure costs

Table 98 Essential training

Technology	NHS staff	Training components	Duration
bkFusion	Urologists, radiologists, radiation oncologists, sonographers, and assisting staff	Not described	One or two days
FusionVu	Urologists, radiologists, nurses, and sonographers	eLearning	2 hours
		On-site training	1 hour
		Live expert support	10 – 15 cases
KOELIS Trinity	End user, consultant, radiologist, CNS	Pre-installation training	3 hours
	OPD staff, theatre staff, ODP	Installation training	1 hour
	End user, consultant, radiologist, CNS	Theatre List	4 or 5 cases

BiopSee	Urologists/radiologists	Not described	3 hours
	Nurses		1 hour
Fusion Bx 2.0	Urologists, nurses and/ or sonographers	Video training	1 hour
		Hands-on training with phantom prostate	0.5-0.75 hours
		Support to clinical cases	10-20 casers over 2-3 days

CNS, clinical nurse specialist; IT, Information Technology; ODP, Operating Department Practitioner; OPD, Outpatient department.

Table 99 Additional time of software fusion vs. cognitive fusion biopsy according to the companies

Fusion system	MRI contouring	Connect fusion system to ultrasound	Contouring ultrasound
bkFusion	3 – 5 minutes	NR	-*
FusionVu	1 minutes	NR	10 seconds
KOELIS Trinity	5 minutes	NR	5 minutes
BiopSee	1-2 minutes	NR	<1 minute
Fusion Bx 2.0	8 -10 minutes	30 seconds	5 – 10 minutes

*Company states that bkFusion does not require ultrasound contouring; NR, not reported

Table 100 Summary of information on the costs of transperineal needle positioning freehand devices in a previous DAR and from the companies' responses to RFIs

Device	Manufacturer	Compatible with	Cost of device	Number of uses	Reprocessing	Co-axial needle	Source
PrecisionPoint	BXTAccelyon	KOELIS Trinity, BiopSee, Fusion Bx 2.0	£206.16	1	-	-	Southampton DAR[126]; Inflated to 2020/2021 price year[163]
			£250.00	NR	NR	NR	KOELIS and Kebomed response to NICE and/or EAG RFI
			£350.00	NR	NR	NR	Focal Healthcare response to NICE and/or EAG RFI
			£150-£250	NR	NR	NR	Medcom response to NICE and/or EAG RFI
FusionVu guide	ExactImaging	FusionVu	£1,333				
EZU-PA3	Hitachi	?	£1971.66**	100***	£5.15	£22.06	Southampton DAR[126]; Inflated to 2020/2021 price year[163]

UA1232	Bk Medical	bkFusion*	£1443.12	100***	£5.15	-	Southampton DAR[126]; Inflated to 2020/2021 price year[163]
Trinity Perine	KOELIS and Kebomed	KOELIS Trinity	£777.64	100	£5.15	-	Southampton DAR[126]; Inflated to 2020/2021 price year[163]
Perine Grid 18G			£779.31	100	NR	used with or without a guide needle	KOELIS and Kebomed response to NICE and/or EAG RFI
Full Grid 18G			£1,303.44	100	NR		
Perine Mini Grid			£86.20	1			
Perine Full Grid			£62.04	1			
SureFire	Delta Surgical	Fusion Bx 2.0	£123.70	1		-	Southampton DAR[126]; Inflated to 2020/2021 price year[163]
			£125.00	NR	NR	NR	Focal Healthcare response to NICE and/or EAG RFI
Unnamed reusable device	NR	BiopSee	£700.00	NR	NR	NR	Medcom response to NICE and/or EAG RFI

*No third-party freehand device validated: **Average unit cost for order of fewer than 5 units (£2000.00) and greater than 5 unit (£1825.50); ***Assumption in Southampton DAR;[126] NR, not reported

Table 101 Disaggregated biopsy costs with LATRUS

LATRUS	bkFusion	FusionVu	KOELIS Trinity	BiopSee	Fusion Bx 2.0	Cognitive fusion
Technology specific						
MRI fusion and US	£57.17	£89.13	£77.44	£46.61	£83.09	£36.15
Installation	£0.01	£0.01	£0.01	£0.01	£0.01	£0.01
Maintenance	£44.65	£37.20	£29.76		£32.32	
Training	£3.43	£0.92	£0.95	£0.68	£0.46	
Procedure time	£22.53	£22.53	£22.53	£22.53	£22.53	
Biopsy setting	£19.67	£19.67	£19.67	£19.67	£19.67	£12.29
TP Biopsy devices						
Total	£147.48	£169.47	£150.37	£89.51	£158.10	£48.44
Not technology specific						
Training	£1.46					
Procedure time	£50.70					
General consumables	£79.10					
Lithomy bed						
Histology	£77.79					
Total	£209.05					
Total per biopsy	£356.53	£378.53	£359.43	£298.56	£367.15	£257.49

Table 102 Disaggregated biopsy costs with LATP

LATP	bkFusion	FusionVu	KOELIS Trinity	BiopSee	Fusion Bx 2.0	Cognitive fusion
Technology specific						
MRI fusion and US	£57.17	£89.13	£78.73	£52.07	£83.52	£36.58
Installation	£0.01	£0.01	£0.01	£0.01	£0.01	
Maintenance	£44.65	£37.20	£29.76		£32.32	
Training	£3.43	£0.92	£0.95	£0.68	£0.46	
Procedure time	£22.53	£22.53	£22.53	£22.53	£22.53	
Biopsy setting	£22.01	£22.01	£22.01	£22.01	£22.01	£14.63
TP Biopsy devices	£81.86	£81.86	£81.86	£81.86	£81.86	£81.86
Total	£231.68	£253.68	£235.87	£179.18	£242.73	£133.07
Not technology specific						
Training	£11.67					
Procedure time	£60.36					
General consumables	£85.44					
Lithomy bed	£3.99					
Histology	£77.79					
Total	£239.25					
Total per biopsy	£470.93	£492.93	£475.12	£418.43	£481.98	£372.32

Table 103 Disaggregated biopsy costs with GATP

GATP	bkFusion	FusionVu	KOELIS Trinity	BiopSee	Fusion Bx 2.0	Cognitive fusion
Technology specific						
MRI fusion and US	£57.17	£89.13	£78.73	£47.42	£83.52	£36.58
Installation	£0.01	£0.01	£0.01	£0.01	£0.01	£0.01
Maintenance	£44.65	£37.20	£29.76		£32.32	
Training	£3.43	£0.92	£0.95	£0.68	£0.46	
Procedure time	£29.83	£29.83	£29.83	£29.83	£29.83	
Biopsy setting	£155.14	£155.14	£155.14	£155.14	£155.14	£132.97
TP Biopsy devices	£90.44	£90.44	£90.44	£90.44	£90.44	£90.44
Total	£380.67	£402.67	£384.86	£323.52	£391.72	£260.00
Not technology specific						
Training	£11.67					
Procedure time	£270.42					
General consumables	£170.29					
Lithomy bed	£3.99					
Histology	£77.79					
Total	£534.15					
Total per biopsy	£914.82	£936.82	£919.01	£857.67	£925.87	£794.15

Table 104 Biopsy procedure adverse event costs

Biopsy adverse events	Cost	Resource use and unit costs
Mild adverse event	£49.78	Resource use for outpatient urinary infection (Wilson (2021)[131]ref, including: - GP visit: £39.23 – PSSRU 2021[163]: General practitioner - unit costs; per patient contact lasting 9.22 minutes - Urinalysis: £10.18 – NHS reference costs 2020/21[166] - Direct Access Pathology Services: currency code DAPS07, Microbiology - 7-day trimethoprim: £0.37 – eMIT 2021[190]- trimethoprim 200 mg x 14 tablets
Non-elective admission*	Transrectal: £2,580.24 Transperineal: £1,952.98	Tamhankar et al. (2020)[186], inflated to 2020/21 price year[163]
Death*	£9,560.56	NHS reference costs 2020/21[166] - Non-Elective: currency code WJ06A, Sepsis with multiple interventions, CC Score 9+ (weighted average of short stay and long stay patients)

*Within 28 days of the procedure; eMIT, electronic market information tool; GP, general practitioner; PSSRU, Personal Social Services Research Unit

Prostate cancer management costs

Table 105 Resource use and costs of monitoring for individuals diagnosed with localised and locally advanced prostate cancer

Treatment assigned	**Active surveillance**				**Radical treatment**					**Resource use and unit costs**
Time	**1st year**			**Subsequent years**	**1st year**			**2nd year**	**Subsequent years**	
Diagnosed CPG	**CPG 1**	**CPG2-3**	**CPG4-5**	**CPG1-5**	**CPG 1**	**CPG2-3**	**CPG4-5**	**CPG1-5**	**CPG1-5**	
Resource use										
PSA test	4			2	2			2	1	£1.85 – NHS reference costs 2020/21[166] – currency code DAPS04, Clinical Biochemistry, Direct Access Pathology Services
Nurse-led outpatient appointment	4			2	2			2	1	£11.00 – assumed as cost per 10 minutes, adjusted from cost per hour of band 7 community-based nurse – PSSRU 2021[163]
DRE	1			1	0			0	0	£78.46 – assumed as cost per approximately 20 minutes of GP appointment – PSSRU 2021[163]; adjusted from General practitioner - unit costs; per patient contact lasting 9.22 minutes
mpMRI	1			0	0			0	0	£294.70 – NHS reference costs 2020/21[166] – currency code RD03Z, Diagnostic Imaging, Magnetic Resonance Imaging Scan of One Area, with Pre- and Post-Contrast
CT scan	0	0.5	0.7	0	0	0.5	0.7	0	0	£150.62 – NHS reference costs 2020/21[166] – currency code RD21A, Diagnostic Imaging, Computerised Tomography Scan of One Area, with Post-Contrast Only, 19 years and over
Bone scan	0	0.5	0.7	0	0	0.5	0.7	0	0	£427.21 – NHS reference costs 2020/21[166] – currency code RN15A, Nuclear Medicine, Nuclear Bone Scan of Two or Three Phases, 19 years and over
Cost per year	£424.56	£713.48	£829.05	£104.16	£25.70	£314.62	£430.18	£25.70	£12.85	

CPG, Cambridge Diagnostic Group; CT, computerized tomography; DRE, digital rectal examination; GP, general practitioner; mpMRI, multiparameter magnetic resonance image; NHS, National Health Service; PSA, prostate-specific antigen; PSSRU, Personal Social Services Research Unit

Table 106 Resource use and costs of radical treatment

Radical treatment	Cost of procedure and follow-up	Resource use and unit costs
Radical prostatectomy	£11,625.37	Robotic surgery: £11,245.08 – NHS reference costs 2020/21[166] – Elective inpatient, currency code LB69Z: Major Robotic, Prostate or Bladder Neck Procedures (Male)
		First surgery appointment: £87.14 – NHS reference costs 2020/21[166] – Outpatient procedure, currency code WF01B: Non-Admitted Face-to-Face Attendance, First (General surgery)
		Two follow-up appointments: 2 x £146.58 – NHS reference costs 2020/21[166] – Outpatient procedure, currency code WF01A: Non-Admitted Face-to-Face Attendance, Follow-up (General surgery)
External radiotherapy	£5,341.81	Preparation: £1,721.79 – NHS reference costs 2020/21[166] Preparation of for Intensity Modulated Radiation Therapy, weighted average of currency codes DC40Z and DC41Z (Total HRGs)
		Fraction delivery – 20 x £181.00 – NHS reference costs 2020/21[166]- Deliver a Fraction of Treatment on a Superficial or Orthovoltage Machine, currency code SC12Z (Total HRGs)
Brachytherapy	£9,156.96	Preparation: £1,550.22– NHS reference costs 2020/21[166]- Preparation for Interstitial Brachytherapy, weighted average of currency code SC55Z over day case, inpatient, outpatient and other setting
		Fraction delivery: £7,606.74 – NHS reference costs 2020/21[166]- Deliver a Fraction of Intraluminal Brachytherapy, weighted average of currency code SC30Z over day case, inpatient outpatient, and other setting

HRG, healthcare resource group; NHS, National Health Service

Table 107 Metastatic treatment costs

Treatment	Cost	Treatments included	Source of unit cost
Metastatic hormone sensitive - year 1	£15,603.87	. ADT: £973.76 for LHRH* (leuprorelin 11.25 mg, every 3 months; triptorelin 11.25mg; or goserelin 3.6mg, every 28 days) + £1.00 one-off bicalutamide 50 mg for 28 days (in year 1 only) . ADT + docetaxel: £973.76 for LHRH* (as above) + £1,404.00 (6 cycles of docetaxel** at a dose of 75 mg/m^2; a cycle every 3 weeks – divided equally over 2 years) + £1.00 one-off bicalutamide 50 mg for 28 days (in year 1 only) . ADT + apalutamide: £973.76 for LHRH (as above) + £35,677.10 (apalatumide 240 mg daily) + £1.00 one-off bicalutamide 50 mg for 28 days (in year 1 only) . ADT + enzalutamide: £973.76 for LHRH* (as above) + £35,672.79 (enzalatumide 160 mg daily) + £1.00 one-off bicalutamide 50 mg for 28 days (in year 1 only)	BNF 2022,[191] eMIT 2022,[190] PSSRU 2021,[163] NHS reference costs 2020/21[166]
Metastatic hormone sensitive - year 2	£15,602.88		
Metastatic hormone resistant - year 1	£14,907.45	. Abiraterone: £23,784.73 (1000 mg daily, 8 months) . Docetaxel**: £4,509.64 (9.5 cycles of docetaxel at a dose of 75 mg/m^2; a cycle every 3 weeks . Enzalutamide: £41,618.26 (160 mg daily, 14 months)	BNF 2022,[191] eMIT 2022,[190] PSSRU 2021,[163] NHS reference costs 2020/21[166]

*Administered by a band 6 hospital-based nurse (15.5 minutes); **Administered by perfusion (NHS reference costs currency codes for delivery of simple parental chemotherapy [SB12Z and SB15Z]); ADT, androgen deprivation therapy; BNF, British national formulary; eMIT, electronic market information tool; LHRH, luteinising hormone-releasing hormone agonists; PSSRU, Personal Social Services Research Unit.

Table 108 Treatment adverse event unit costs

Treatment for	Adverse event	Unit cost	Source
Localised and locally advanced prostate cancer	Erectile dysfunction	£328.58	NHS reference costs 2020/21[166] – Treatment of Erectile Dysfunction weighted average of the currency code LB43Z General Surgery, Genitourinary Medicine, Plastic Surgery, Urology
	Urinary incontinence	£317.54	NICE NG 131[133] – managed by containment pads. Inflated to 2020/2021 price year[163]
	Bowel dysfunction	£1,941.19	NICE NG 131[133] – mean weighted cost including costs associated with sigmoidoscopy, laser therapy, enemas and blood transfusion. Inflated to 2020/2021 price year[163]
Hormone sensitive metastatic prostate cancer	Blood disorder	£2,428.70	NHS reference costs 2020/21[166] – weighted average of currency codes SA03G-SA03H, SA08G-SA08J, SA12G-SA12K non-elective long stay and non-elective short stay
	Cardiac disorder	£2,042.04	NHS reference costs 2020/21[166] – weighted average of currency codes EB10A-EB10E non-elective long stay and non-elective short stay
	Endocrine disorder	£328.58	Assume the same as erectile dysfunction (as in Southampton DAR)[126]
	Gastrointestinal disorder	£2,019.47	NHS reference costs 2020/21[166] – weighted average of currency codes FD10A-FD10M non-elective long stay and non-elective short stay
	General disorder	£39.90	. GP visit per patient contact lasting 9.22 minutes: £39.23 – General practitioner - unit costs; PSSRU 2021[163] . 3-day Trimethoprim: £0.67 - eMIT 2021[190] - trimethoprim 200 mg x 6 tablets
	Musculoskeletal disorder	£26.58	NHS reference costs 2020/21[166] – weighted average of currency codes HD26D-HD26G non-elective long stay and non-elective short stay
	Nervous system disorder	£1,933.29	NHS reference costs 2020/21[166] – weighted average of currency codes AA26C-AA26H non-elective long stay and non-elective short stay
	Neutropenia	£9,842.93	NHS reference costs 2020/21[166] – weighted average of currency codes PM45A-PM45D non-elective long stay and non-elective short stay
	Renal disorder	£49.78	Assume the same as urinary infection (as in Southampton DAR)[126]

	Respiratory disorder	£971.68	NHS reference costs 2020/21[166] – weighted average of currency codes DZ19H-DZ19N non-elective long stay and non-elective short stay
	Skin disorder	£2,191.91	NHS reference costs 2020/21[166] – weighted average of currency codes JD07A-JD07K non-elective long stay and non-elective short stay

DAR, diagnostic assessment report; eMIT, electronic market information tool; GP, general practitioner; NHS, National Health Service; NICE, National Institute of Health and Care Excellence; PSSRU, Personal Social Services Research Unit.

Table 109 Model parameters

Parameter	Value		Probabilistic setup	Source
Population characteristics				
. Age	66 years		NA	Southampton DAR[126]
. Prevalence & distribution across ISUP grade				
.. No PCa	0.12		Calculated from each 1,000 iterations of network 1 and 2	See Section 6.3.1
.. ISUP grade 1	0.32			
.. ISUP grade 2	0.26			
.. ISUP grade 3	0.18			
.. ISUP grade 4-5	0.12			
Diagnostic performance				
. 1st biopsy and repeat biopsy with cognitive fusion				
Probability of [diagnosis] | [true disease]	Targeted	Combined		
.. ISUP grade 4-5 | ISUP grade 4-5	0.552	0.573	Calculated from each 1,000 iterations of network 1 for targeted and network 2 for combined	See Section 6.3.1
.. ISUP grade 3 | ISUP grade 4-5	0.101	0.140		
.. ISUP grade 2 | ISUP grade 4-5	0.111	0.130		
.. ISUP grade 1 | ISUP grade 4-5	0.111	0.047		
.. No PCa | ISUP grade 4-5	0.125	0.111		
.. ISUP grade 3 | ISUP grade 3	0.479	0.510		
.. ISUP grade 2 | ISUP grade 3	0.192	0.207		
.. ISUP grade 1 | ISUP grade 3	0.140	0.059		
.. No PCa | ISUP grade 3	0.189	0.224		
.. ISUP grade 2 | ISUP grade 2	0.338	0.544		
.. ISUP grade 1 | ISUP grade 2	0.362	0.204		
.. No PCa | ISUP grade 2	0.300	0.251		
.. ISUP grade 1 | ISUP grade 1	0.171	0.329		
.. No PCa | ISUP grade 1	0.829	0.671		
.. No PCa | No PCa	1.000	1.000		
. 1st biopsy and repeat biopsy with software fusion				
Probability of [diagnosis] | [true disease]	Targeted	Combined		
.. ISUP grade 4-5 | ISUP grade 4-5	0.281	0.724	Calculated from each 1,000 iterations from network 1 for targeted	See Section 6.3.1
.. ISUP grade 3 | ISUP grade 4-5	0.163	0.071		

.. ISUP grade 2 | ISUP grade 4-5	0.173	0.066	and network 2 for combined	
.. ISUP grade 1 | ISUP grade 4-5	0.187	0.070		
.. No PCa | ISUP grade 4-5	0.195	0.069		
.. ISUP grade 3 | ISUP grade 3	0.616	0.603		
.. ISUP grade 2 | ISUP grade 3	0.134	0.130		
.. ISUP grade 1 | ISUP grade 3	0.124	0.135		
.. No PCa | ISUP grade 3	0.126	0.132		
.. ISUP grade 2 | ISUP grade 2	0.314	0.770		
.. ISUP grade 1 | ISUP grade 2	0.437	0.152		
.. No PCa | ISUP grade 2	0.249	0.078		
.. ISUP grade 1 | ISUP grade 1	0.291	0.472		
.. No PCa | ISUP grade 1	0.709	0.528		
.. No PCa | No PCa	1.000	1.000		
Probability of repeat biopsy				
. if diagnosed as No PCa	5%		NA	Southampton DAR assumption[126]
. if diagnosed as ISUP grade 1	15.45%		Beta distribution: α=95; β=520	Southampton DAR[126]
Biopsy adverse events rates				
. Mild adverse events with TR biopsy	1.31%		Beta distribution: α=15; β=1132	Southampton DAR[126, 160]
. Mild adverse events with TP biopsy	9.13%		Beta distribution: α=274; β=2726	
. Leading to NEL with TR biopsy	3.74%		Beta distribution: α=2845; β=73261	
. Leading to NEL with TR biopsy	3.54%		Beta distribution: α=1314; β=35763	
. TR mortality	0.07%		Beta distribution: α=53; β=76053	
. TP mortality	0.05%		Beta distribution: α=19; β=37058	
Distribution by biopsy approach at 1st biopsy				
. LATRUS	35%		NA	Assumption informed by NHS reference data 2018/19[187]
. LATP	65%		NA	
Distribution by biopsy approach at repeat biopsy				
. LATRUS	30%		NA	Assumption informed by NHS reference data 2018/19[187] and clinical advice
. LATP	60%		NA	
. GATP	10%		NA	
Long-term model transitions				

. Progression Localised/Locally advanced to Metastatic			
.. Lambda CPG 1 with observation	0.0143	Sampled from 1000 simulations of the calibration model joint output for the 4 CPG categories and treatment received	Calibrated see Section 6.3.3
.. Lambda CPG 2 with observation	0.0379		
.. Lambda CPG 3 with observation	0.1197		
.. Lambda CPG 4-5 with observation	0.3997		
.. Lambda CPG 1 with radical treatment	0.0063		
.. Lambda CPG 2 with radical treatment	0.0164		
.. Lambda CPG 3 with radical treatment	0.0514		
.. Lambda CPG 4-5 with radical treatment	0.1683		
. Metastatic to PCa death			
Weibull	γ= 1.26; λ=0.11	Multivariate lognormal	The PCa death curve for the control arm in Clarke (2019)[151] was digitised by using WebPlotDigitizer[153]; a pseudo-IPD was reconstructed by using Guyot algorithm[154], Weibull distribution was then fitted to the pseudo-IPD to obtain γ, λ and variance-covariance matrix using *flexsurv* package in R[155] See Section 6.3.3
.. Mortality HR for Docetaxel +ADT vs ADT alone	0.78	Log-normal, 95% CI (0.66-0.93)	James et al. (2016)[59]
.. Mortality HR for Enzalutamide +ADT vs ADT alone	0.66	Log-normal, 95% CI (0.53-0.81)	ARCHES study[157]
.. Mortality HR for Apalutamide +ADT vs ADT alone	0.65	Log-normal, 95% CI (0.53-0.79)	TITAN study[158]
. Other cause mortality	Age dependent	NA	ONS lifetables 2018-2020[152]
Treatment distributions			
. Localised disease/Locally advanced disease			
.. Radical treatment and diagnosed (ISUP grade 4-5)	75.9%	Dirichlet distribution	Calculated as sum of proportions of radical prostatectomy and radiotherapy; Parry et al. (2020)[159]
.. Radical treatment and diagnosed (ISUP grade 3)	66.3%		
.. Radical treatment and diagnosed (ISUP grade 2)	48.4%		
.. Radical treatment and diagnosed (ISUP grade 1)	11.3%		
.. Radical treatment and diagnosed (No PCa)	0%	NA	Assumption
. Metastatic cancer			

.. ADT	50.0%	NA	Assumption informed by Southampton DAR[126] and NPCA report 2021[188]
.. ADT + docetaxel	9.4%	NA	
.. ADT + apalutamide	6.6%	NA	
.. ADT + enzalutamide	34.0%	NA	
Treatment adverse event rates			
. Radical prostatectomy			
.. Sexual dysfunction	85.39%	Beta distribution: α=304; β=52	Southampton DAR [126]
.. Bowel dysfunction	2.47%	Beta distribution: α=9; β=355	
.. Urinary dysfunction	26.24%	Beta distribution: α=95; β=267	
. Radiotherapy			
.. Sexual dysfunction	62.39%	Beta distribution: α=219; β=132	Southampton DAR [126]
.. Bowel dysfunction	5.85%	Beta distribution: α=21; β=338	
.. Urinary dysfunction	3.63%	Beta distribution: α=21; β=345	
. Active surveillance			
.. Erectile dysfunction	50.88%	Beta distribution: α=173; β=167	Southampton DAR [126]
.. Bowel dysfunction	1.68%	Beta distribution: α=6; β=352	
.. Urinary incontinence	4.20%	Beta distribution: α=15; β=342	
. Metastatic treatment			
.. ADT			
… Blood disorder	0.00%		Southampton DAR [126]
… Cardiac disorder	2.96%	Beta distribution: α=35; β=1149	
… Endocrine disorder	12.25%	Beta distribution: α=145; β=1039	
… Gastrointestinal disorder	3.04%	Beta distribution: α=36; β=1148	
… General disorder	3.89%	Beta distribution: α=46; β=1138	
… Musculoskeletal disorder	5.83%	Beta distribution: α=69; β=1115	
… Nervous system disorder	1.69%	Beta distribution: α=20; β=1164	
… Neutropenia	1.77%	Beta distribution: α=21; β=1163	
… Renal disorder	6.00%	Beta distribution: α=71; β=1113	
… Respiratory disorder	2.28%	Beta distribution: α=27; β=1157	
… Skin disorder	0.00%		
.. ADT + Docetaxel			
… Blood disorder	0.00%		Southampton DAR [126]

… Cardiac disorder	2.91%	Beta distribution: α=16; β=534	
… Endocrine disorder	10.36%	Beta distribution: α=57; β=493	
… Gastrointestinal disorder	8.18%	Beta distribution: α=45; β=505	
… General disorder	6.18%	Beta distribution: α=34; β=516	
… Musculoskeletal disorder	5.82%	Beta distribution: α=32; β=518	
… Nervous system disorder	3.45%	Beta distribution: α=19; β=531	
… Neutropenia	27.27%	Beta distribution: α=150; β=400	
… Renal disorder	4.18%	Beta distribution: α=23; β=527	
… Respiratory disorder	5.27%	Beta distribution: α=29; β=521	
… Skin disorder	0.00%		
.. ADT + Apalutamide			
… Blood disorder	2.10%	Beta distribution: α=11; β=513	Southampton DAR [126]
… Cardiac disorder	8.40%	Beta distribution: α=44; β=480	
… Endocrine disorder	0.00%		
… Gastrointestinal disorder	1.15%	Beta distribution: α=6; β=518	
… General disorder	3.44%	Beta distribution: α=18; β=506	
… Musculoskeletal disorder	6.49%	Beta distribution: α=34; β=490	
… Nervous system disorder	0.19%	Beta distribution: α=1; β=523	
… Neutropenia	0.00%		
… Renal disorder	0.76%	Beta distribution: α=4; β=520	
… Respiratory disorder	0.00%		
… Skin disorder	6.49%	Beta distribution: α=34; β=490	
.. ADT + Enzalutamide			
… Blood disorder	0.00%		Southampton DAR [126]
… Cardiac disorder	4.90%	Beta distribution: α=28; β=544	
… Endocrine disorder	0.35%	Beta distribution: α=2; β=570	
… Gastrointestinal disorder	0.52%	Beta distribution: α=3; β=569	
… General disorder	2.80%	Beta distribution: α=16; β=556	
… Musculoskeletal disorder	4.37%	Beta distribution: α=25; β=547	

… Nervous system disorder	2.10%	Beta distribution: α=12; β=560	
… Neutropenia	0.35%	Beta distribution: α=2; β=570	
… Renal disorder	0.00%		
… Respiratory disorder	0.00%		
… Skin disorder	0.35%	Beta distribution: α=2; β=570	
HRQoL			
. Disutility of biopsy adverse events			
.. Mild adverse events	-0.289	NA	Southampton DAR;[126] assumed duration 3 days
.. Leading to NEL	-0.490	NA	Southampton DAR;[126] assumed duration 30 days
.. Death	-0.490	NA	Southampton DAR;[126] assumed duration 30 days
. Baseline health state utility	Age and sex dependent	NA	Ara and Brazier (2010)[147]
. Localised treatment disutility			
.. Sexual dysfunction	-0.0230	No-mild symptoms: Beta distribution: α=578; β=93 Moderate-severe symptoms: Beta distribution: α=452; β=87	Calculated as the difference between no-mild symptoms and moderate-severe symptoms (as per Southampton DAR)[126]
.. Urinary dysfunction	-0.0950	No-mild symptoms: Beta distribution: α=1013; β=154 Moderate-severe symptoms: Beta distribution: α=131; β=39	
.. Bowel dysfunction	-0.2090	No-mild symptoms: Beta distribution: α=1097; β=176 Moderate-severe symptoms: Beta distribution: α=62; β=33	
. Metastatic disutility	- 0.137	Localised 1: Beta distribution: α=102; β=11 Localised 2: Beta distribution: α=404; β=50 Localised 3: Beta distribution: α=841; β=126 Metastatic: Beta distribution: α=165; β=58	Calculated as the difference between metastatic and the average across localised 1, 2, 3 (as per Southampton DAR)[126]
Resource use and costs			

. Annual patient throughput	300	NA	Assumed based on NHS reference costs 2018/19[187]
. Cost per first cognitive fusion biopsy (targeted or combined)	£332.13	NA	Calculated
. Cost per first software fusion (targeted or combined)	£427.33	NA	Calculated
. Cost per repeat cognitive fusion biopsy (targeted or combined)	£380.05	NA	Calculated
. Cost per repeat software fusion (targeted or combined)	£477.42	NA	Calculated
. Cost of localised treatment			
.. Cost of radical prostatectomy	£11,625.37	NA	Calculated
.. Cost of radiotherapy for those who diagnosed as CPG1	£6,283.42	NA	Calculated
.. Cost of radiotherapy for those who diagnosed as CPG2	£5,754.11	NA	Calculated
.. Cost of radiotherapy for those who diagnosed as CPG3	£5,510.29	NA	Calculated
.. Cost of radiotherapy for those who diagnosed as CPG4-5	£5,402.04	NA	Calculated
.. Cost of ADT	£973.76	NA	Calculated, see Table 107
.. Cost of bicalutamide	£1.49	NA	21 days course of bicalutamide - BNF 2022[191] bicalutamide 50mg x 28 tablets
. Cost of metastatic treatment			
.. Cost of 1st year hormone-sensitive treatment	£15,603.87	NA	Calculated, see Table 107
.. Cost of 2nd year hormone-sensitive treatment	£15,602.88	NA	
.. Cost of metastatic treatment in subsequent years (one-off)	£14,907.45	NA	
. Cost of monitoring/ active surveillance			
.. Cost of AS for those who diagnosed as CPG1 in 1st year	£424.56	NA	Calculated
.. Cost of AS for those who diagnosed as CPG2-3 in 1st year	£713.48	NA	Calculated
.. Cost of AS for those who diagnosed as CPG4-5 in 1st year	£829.05	NA	Calculated

.. Cost of AS for those who diagnosed as any CPG in subsequent years	£104.16	NA	Calculated
.. Cost of monitoring those who diagnosed as CPG1 receiving RT, in 1st year	£25.70	NA	Calculated
.. Cost of monitoring those who diagnosed as CPG2-3 receiving RT, in 1st year	£314.62	NA	Calculated
.. Cost of monitoring those who diagnosed as CPG4-5 receiving RT, in 1st year	£430.18	NA	Calculated
.. Cost of monitoring those who diagnosed as any CPG receiving RT, in 2nd year	£ 25.70	NA	Calculated
.. Cost of monitoring those who diagnosed as any CPG receiving RT, in 2+ year	£12.85	NA	Calculated
.. Cost of monitoring those who have No PCa diagnosed as No PCa	£158.99	NA	Calculated
.. Cost of monitoring those who have No PCa diagnosed as ISUP grade 1	£242.02	NA	Calculated
.. Cost of monitoring metastatic patients (one off)	£577.83	NA	Calculated
. Cost of managing adverse events			
.. Cost of managing adverse events of biopsy procedure			
… Cost per mild adverse event	£49.78	NA	Calculated
… Cost per NEL event with LATRUS	£2,580.24	NA	Calculated
… Cost per NEL event with LATP/GATP	£1,952.98	NA	Calculated
… Cost per biopsy death	£9,560.56	NA	Calculated
.. Cost of managing adverse events of			
… Active surveillance	£213.06	NA	Calculated
… Radical prostatectomy	£411.91	NA	Calculated
… Radiotherapy	£330.09	NA	Calculated
… Metastatic treatment			
…. ADT	£397.49	See probabilistic setup for AE rates of ADT	Calculated
…. ADT+Docetaxel	£3,067.24	See probabilistic setup for AE rates of ADT+ Docetaxel	Calculated

…. ADT+Enzalutamide	£196.62	See probabilistic setup for AE rates of ADT+Enzalutamide	Calculated
…. ADT+Apalutamide	£394.96	See probabilistic setup for AE rates of ADT+Apalutamide	Calculated
. End of life costs	£16,546.08	NA	Round (2015)[149]; inflated to 2020/2021 price year[163]

|, conditional on; ADT, androgen deprivation therapy; AE, adverse event; AS, active surveillance; CPG, Cambridge Prognostic Group; GATP, general anaesthesia transperineal; HR, hazard ratio; IPD, individual patient data; ISUP, International Society of Urological Pathology; LATP, local anaesthesia transperineal; LATRUS, local anaesthesia transrectal; NA, not applicable; NA, not applicable; NEL, non-elective admission; ONS, Office for National Statistics; PCa, prostate cancer; RT, radical treatment; TP, transperineal; TR, transrectal.

APPENDIX 12. ADDITIONAL COST EFFECTIVENESS RESULTS

Base-case analysis

Table 110 Deterministic base-case prevalence, and final classification from the diagnostic pathway: targeted biopsy

Strategy	Prevalence					Proportion correctly classified					
	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	All categories
Cognitive fusion	0.116	0.183	0.262	0.318	0.121	0.066	0.090	0.095	0.057	0.121	0.428
Software fusion						0.067	0.095	0.149	0.108	0.121	0.540

*Final classification in the model

Table 111 Deterministic base-case diagnostic pathway events, and disaggregated costs and QALY loss: targeted biopsy

Strategy	Proportion repeat biopsy		Proportion adverse events			Cost			AEs QALY loss
	All	Unnecessary*	Death	Mild	repeat biopsy	1st biopsy	Repeat biopsy	AEs	
Cognitive fusion	0.055	0.038	0.001	0.068	0.038	£332	£21	£92	-0.00176
Software fusion	0.050	0.035	0.001	0.067	0.038	£427	£24	£92	-0.00175

*Unnecessary biopsy is defined as a second biopsy that did not raise the ISUP grade to at least 2; AEs, adverse events

Table 112 Deterministic base-case long-term undiscounted disaggregated costs: targeted biopsy

Strategy	Local disease – Radical treatment					Local disease adverse events					Metastatic disease		Monitoring	EoL
	Immediate	Delayed				Immediate	Delayed							
	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	Treatment	AEs		
Cognitive fusion	£1,844	£103	£280	£568	£252	£688	£12	£84	£276	£146	£17,241	£456	£948	£16,510

Software fusion	£2,158	£76	£260	£395	£202	£850	£10	£78	£192	£117	£17,008	£449	£1,047	£16,510

AEs, adverse events; EoL, end of life

Table 113 Deterministic base-case long-term undiscounted disaggregated health outcomes: targeted biopsy

Strategy	LYs	Baseline QALYs	QALY loss		
			Immediate radical treatment	Delayed radical treatment	Metastatic disease
Cognitive fusion	16.22	10.99	-0.09	-0.13	-0.52
Software fusion	16.25	11.01	-0.13	-0.10	-0.51

LYs, life-years

Table 114 Deterministic base-case prevalence, and final classification from the diagnostic pathway: combined biopsy

Strategy	Prevalence					Proportion correctly classified					
	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	All categories
Cognitive fusion	0.116	0.183	0.262	0.318	0.121	0.034	0.115	0.089	0.096	0.121	0.455
Software fusion						0.085	0.113	0.207	0.154	0.121	0.680

*Final classification in the model

Table 115 Deterministic base-case diagnostic pathway events, and disaggregated costs and QALY loss: combined biopsy

Strategy	Proportion repeat biopsy		Proportion adverse events			Cost			AEs QALY loss
	All	Unnecessary*	Death	Mild	repeat biopsy	1st biopsy	Repeat biopsy	AEs	
Cognitive fusion	0.062	0.043	0.001	0.068	0.038	£332	£23	£93	-0.00177
Software fusion	0.051	0.036	0.001	0.067	0.038	£427	£25	£92	-0.00176

*Unnecessary biopsy is defined as a second biopsy that did not raise the ISUP grade to at least 2; AEs, adverse events

Table 116 Deterministic base-case long-term undiscounted disaggregated costs: combined biopsy

	Local disease – Radical treatment					Local disease adverse events					Metastatic disease		Monitoring	EoL
	Immediate	Delayed				Immediate	Delayed							
Strategy	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	Treatment	AEs		
Cognitive fusion	£1,835	£169	£205	£574	£214	£723	£20	£62	£280	£124	£17,172	£454	£1,016	£16,510
Software fusion	£2,547	£57	£216	£181	£158	£1,048	£7	£65	£88	£92	£16,705	£441	£1,167	£16,510

AEs, adverse events; EoL, end of life

Table 117 Deterministic base-case long-term undiscounted disaggregated health outcomes: combined biopsy

Strategy	LYs	Baseline QALYs	QALY loss		
			Immediate radical treatment	Delayed radical treatment	Metastatic disease
Cognitive fusion	16.21	10.98	-0.11	-0.12	-0.52
Software fusion	16.29	11.03	-0.18	-0.07	-0.50

LYs, life-years

Table 118 Deterministic base-case cost-effectiveness results: combined software fusion technologies pairwise comparisons wih cognitive fusion

Strategy	Diagnostic model	Overall results					
	Inc costs	Total LYs*	Total QALYs*	Total Costs*	ICER vs. cognitive fusion**	NHB at £20,000**	NHB at £30,000**
Targeted cognitive fusion	-	11.75	8.68	£22,457	-	7.56	7.93

Combined software fusion	£99	11.76	8.69	£22,536	£9,285	7.56	7.94
Combined bkFusion	£103			£22,540	£9,725	7.56	7.94
Combined FusionVu	£126			£22,563	£12,443	7.56	7.94
Combined Koelis Trinity	£106			£22,544	£10,187	7.56	7.94
Combined Fusion Bx 2.0	£114			£22,551	£11,072	7.56	7.94
Combined BiopSee	£45			£22,483	£2,998	7.57	7.94

Figure 17 Targeted software fusion cost threshold analysis

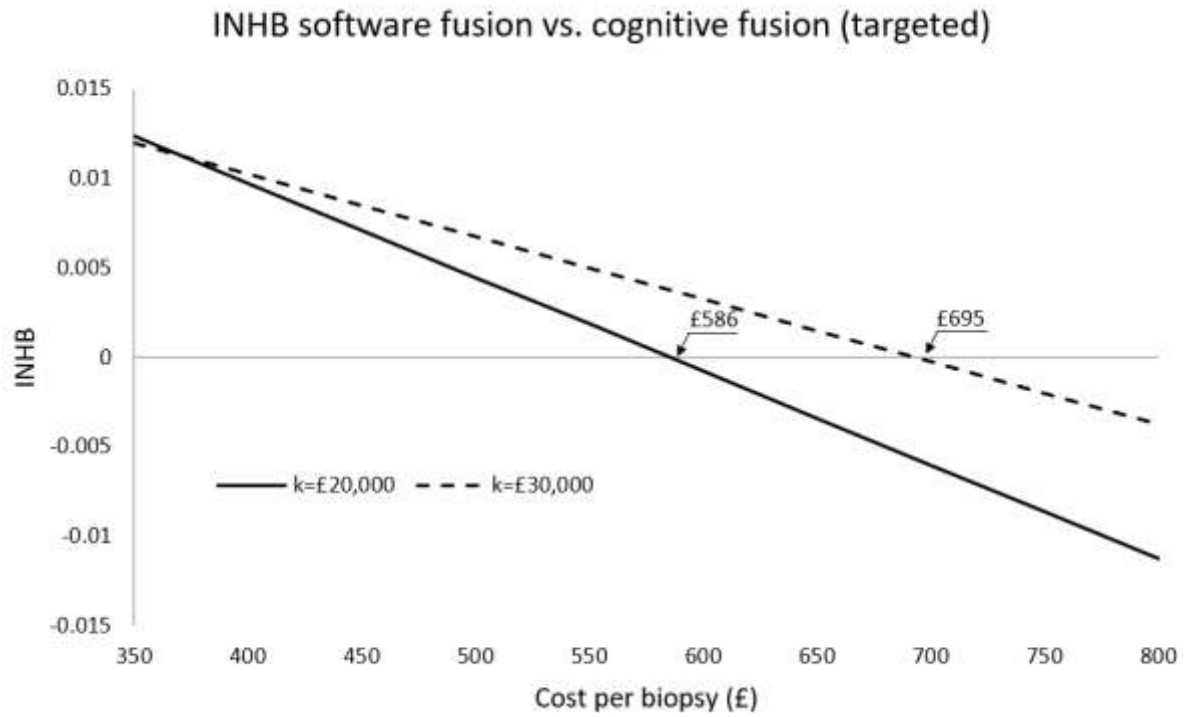
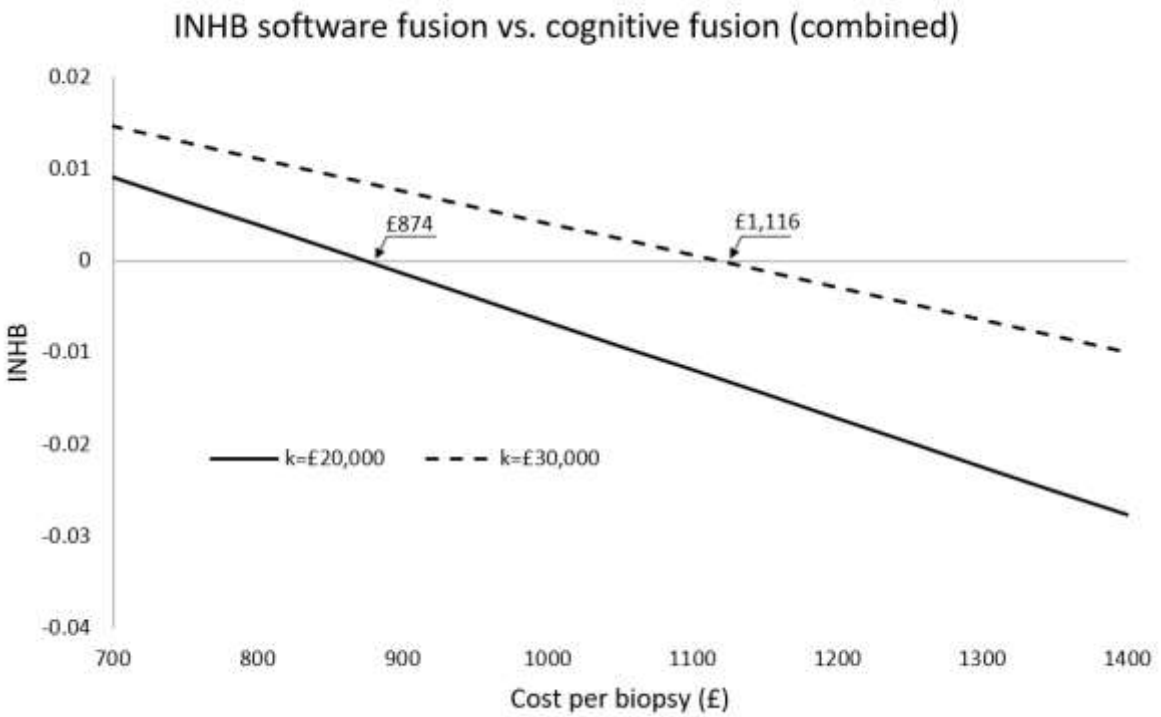


Figure 18 Software fusion cost threshold analysis



Subgroup analyses

Table 119 Deterministic results for prior biopsy subgroup prevalence, and final classification from the diagnostic pathway: targeted biopsy

	Prevalence					Proportion correctly classified					
Strategy	**CPG 4-5**	**CPG G3**	**CPG 2**	**CPG 1**	**No PCa**	**CPG 4-5**	**CPG G3**	**CPG 2**	**CPG 1**	**No PCa**	**All categories**
Cognitive fusion	0.085	0.131	0.132	0.224	0.428	0.048	0.063	0.043	0.033	0.428	0.614
Software fusion						0.050	0.070	0.074	0.070	0.428	0.692

*Final classification in the model

Table 120 Deterministic for prior biopsy subgroup analysis diagnostic pathway events, and disaggregated costs and QALY loss: targeted biopsy

Strategy	Proportion repeat biopsy		Proportion adverse events			Cost			AEs QALY loss
	All	Unnecessary*	Death	Mild	NEL	1st biopsy	Repeat biopsy	AEs	
Cognitive fusion	0.052	0.042	0.001	0.067	0.038	£332	£20	£92	-0.00176
Software fusion	0.049	0.040	0.001	0.067	0.038	£427	£23	£92	-0.00175

*Unnecessary biopsy is defined as a second biopsy that did not raise the ISUP grade to at least 2; AEs, adverse events, NEL, leading to non-elective admissions

Table 121 Deterministic for prior biopsy subgroup long-term undiscounted disaggregated costs: targeted biopsy

Strategy	Local disease – Radical treatment					Local disease adverse events					Metastatic disease		Monitoring	EoL
	Immediate	Delayed				Immediate	Delayed							
	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	Treatment	AEs		
Cognitive fusion	£1,185	£78	£202	£303	£183	£416	£9	£61	£148	£107	£11,439	£302	£1,109	£16,509
Software fusion	£1,394	£55	£183	£204	£148	£522	£7	£55	£99	£86	£11,287	£298	£1,177	£16,509

AEs, adverse events; EoL, end of life

Table 122 Deterministic for prior biopsy subgroup long-term undiscounted disaggregated health outcomes: targeted biopsy

Strategy	LYs	Baseline QALYs	QALY loss		
			Immediate radical treatment	Delayed radical treatment	Metastatic disease
Cognitive fusion	16.72	11.27	-0.05	-0.09	-0.35
Software fusion	16.74	11.28	-0.08	-0.06	-0.34

LYs, life-years

Table 123 Deterministic results for prior biopsy subgroup prevalence, and final classification from the diagnostic pathway: combined biopsy

Strategy	Prevalence					Proportion correctly classified					
	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	All categories
Cognitive fusion	0.085	0.131	0.132	0.224	0.428	0.028	0.081	0.040	0.071	0.428	0.648
Software fusion						0.060	0.081	0.100	0.115	0.428	0.784

*Final classification in the model

Table 124 Deterministic results for prior biopsy subgroup diagnostic pathway events, and disaggregated costs and QALY loss: combined biopsy

Strategy	Proportion repeat biopsy		Proportion adverse events			Cost			AEs QALY loss
	All	Unnecessary*	Death	Mild	NEL	1st biopsy	Repeat biopsy	AEs	
Cognitive fusion	0.057	0.046	0.001	0.068	0.038	£332	£22	£92	-0.00177
Software fusion	0.053	0.043	0.001	0.068	0.038	£427	£25	£92	-0.00176

*Unnecessary biopsy is defined as a second biopsy that did not raise the ISUP grade to at least 2; AEs, adverse events; NEL, leading to non-elective admissions

Table 125 Deterministic for prior biopsy subgroup long-term undiscounted disaggregated costs: combined biopsy

Strategy	Local disease – Radical treatment					Local disease adverse events					Metastatic disease		Monitoring	EoL
	Immediate	Delayed				Immediate	Delayed							
	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	Treatment	AEs		
Cognitive fusion	£1,187	£124	£152	£306	£147	£448	£14	£46	£149	£85	£11,385	£301	£1,172	£16,508
Software fusion	£1,612	£45	£157	£105	£104	£633	£6	£47	£51	£61	£11,127	£294	£1,272	£16,508

AEs, adverse events; EoL, end of life

Table 126 Deterministic for prior biopsy subgroup long-term undiscounted disaggregated health outcomes: combined biopsy

Strategy	LYs	Baseline QALYs	QALY loss		
			Immediate radical treatment	Delayed radical treatment	Metastatic disease
Cognitive fusion	16.71	11.27	-0.07	-0.08	-0.34
Software fusion	16.76	11.29	-0.11	-0.05	-0.34

LYs, life years; QALY, quality adjusted life-years

Scenario analyses

Table 127 Deterministic results for scenario 1 - PAIREDCAP (2019) baseline - cost-effectiveness results: targeted biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00175	£442	11.17	7.99	£32,490	11.17	7.99	£32,932		6.34	6.89
Software fusion	-0.00174	£538	11.19	8.00	£32,432	11.19	7.99	£32,970		6.35	6.90
	Inc QALY loss	Inc Costs	Inc LYs*	Inc QALYs*	Inc Costs*	Inc LYs*	Inc QALYs*	Inc Costs*		INHB at £20,000**	INHB at £30,000**
Software fusion vs. cognitive fusion	0.00001	£96	0.02	0.01	-£58	0.02	0.01	£39	£4,428	0.01	0.01

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 128 Deterministic results for scenario - PAIREDCAP (2019) baseline - prevalence, and final classification from the diagnostic pathway: targeted biopsy

	Prevalence					Proportion correctly classified					
Strategy	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	CPG 4-5	CPG G3	CPG 2	CPG 1	No PCa	All categories
Cognitive fusion	0.169	0.252	0.322	0.226	0.031	0.103	0.135	0.140	0.058	0.031	0.467
Software fusion						0.100	0.136	0.193	0.086	0.031	0.544

*Final classification in the model

Table 129 Deterministic results for scenario 1 - PAIREDCAP (2019) baseline - diagnostic pathway events, and disaggregated costs and QALY loss: targeted biopsy

	Proportion repeat biopsy		Proportion adverse events			Cost			AEs QALY loss
Strategy	All	Unnecessary*	Death	Mild	NEL	1st biopsy	Repeat biopsy	AEs	
Cognitive fusion	0.048	0.028	0.001	0.067	0.038	£332	£18	£92	-0.00175
Software fusion	0.042	0.025	0.001	0.067	0.038	£427	£20	£91	-0.00174

*Unnecessary biopsy is defined as a second biopsy that did not raise the ISUP grade to at least 2; AEs, adverse events; NEL, leading to non-elective hospital admission

Table 130 Deterministic results for scenario 1 - PAIREDCAP (2019) baseline - long-term undiscounted disaggregated costs: targeted biopsy

	Local disease – Radical treatment					Local disease adverse events					Metastatic disease		Monitoring	EoL
	Immediate	Delayed				Immediate	Delayed							
Strategy	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	All CPG	CPG 4-5	CPG 3	CPG 2	CPG 1	Treatment	AEs		
Cognitive fusion	£2,638	£134	£350	£626	£162	£970	£16	£105	£304	£94	£21,381	£565	£934	£16,512
Software fusion	£2,937	£107	£325	£449	£136	£1,120	£14	£98	£218	£79	£21,154	£559	£998	£16,512

AEs, adverse events; EoL, end of life

Table 131 Deterministic results for scenario 1 - PAIREDCAP (2019) baseline - long-term undiscounted disaggregated health outcomes: targeted biopsy

Strategy	LYs	Baseline QALYs	QALY loss		
			Immediate radical treatment	Delayed radical treatment	Metastatic disease
Cognitive fusion	15.78	10.73	-0.13	-0.10	-0.65
Software fusion	15.80	10.75	-0.16	-0.09	-0.64

LYs, life-years

Table 132 Deterministic results for scenario 2 - Zhou (2018) diagnostic - cost-effectiveness results: targeted biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****
Cognitive fusion	-0.00176	£446	11.55	8.41	£26,652	11.55	8.40	£27,098		7.05	7.50
Software fusion	-0.00175	£543	11.58	8.43	£26,638	11.58	8.43	£27,180		7.07	7.52
	Inc QALY loss	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00001	£97	0.03	0.03	-£14	0.03	0.03	£83	£3,105	0.02	0.02

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 133 Deterministic results for scenario 3 - degradation of repeat biopsy accuracy - cost-effectiveness results: targeted biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****

Cognitive fusion	-0.00176	£445	11.44	8.29	£27,922	11.44	8.29	£28,367		6.87	7.34
Software fusion	-0.00175	£543	11.46	8.30	£27,887	11.46	8.30	£28,429		6.88	7.35
	Inc QALY loss	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00001	£98	0.02	0.01	-£35	0.02	0.01	£63	£5,477	0.01	0.01

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 134 Deterministic results for scenario 3 - degradation of repeat biopsy accuracy - cost-effectiveness results: combined biopsy

Strategy	**Diagnostic model**		**Long-term model**			**Overall results**					
	QALY loss	**Total Costs**	**Total LYs***	**Total QALYs***	**Total Costs***	**Total LYs***	**Total QALYs***	**Total Costs***	**ICER****	**NHB at £20,000****	**NHB at £30,000****
Cognitive fusion	-0.00177	£448	11.44	8.28	£27,892	11.44	8.28	£28,340		6.86	7.33
Software fusion	-0.00176	£544	11.49	8.31	£27,843	11.49	8.30	£28,386		6.88	7.36
	Inc QALY loss	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00002	£95	0.05	0.03	-£49	0.05	0.03	£46	£1,801	0.02	0.02

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 135 Deterministic results for scenario 4 -software fusion as quality assurance - cost-effectiveness results: targeted biopsy

	Diagnostic model	**Long-term model**	**Overall results**

Strategy	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00176	£445	11.45	8.29	£27,864	11.45	8.29	£28,310		6.87	7.34
Software fusion	-0.00174	£537	11.45	8.29	£27,859	11.45	8.29	£28,396		6.87	7.34
	Inc QALY loss	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00002	£92	0.00	0.00	-£6	0.00	0.00	£87	£874,744	0.00	0.00

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 136 Deterministic results for scenario 4 -software fusion as quality assurance - cost-effectiveness results: combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00177	£448	11.44	8.28	£27,833	11.44	8.28	£28,282		6.86	7.34
Software fusion	-0.00174	£538	11.44	8.28	£27,824	11.44	8.28	£28,363		6.86	7.33
	Inc QALY loss	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00003	£90	0.00	0.00	-£9	0.00	0.00	£81	£581,847	0.00	0.00

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 137 Deterministic results for scenario 5 - radical treatment for all identified CPG≥2 and conservative treatment for CPG1 - cost-effectiveness results: targeted biopsy

Strategy	Diagnostic model		Long-term model			Overall results			ICER**	NHB at £20,000**	NHB at £30,000**
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*			
Cognitive fusion	-0.00176	£445	11.55	8.37	£28,816	11.55	8.37	£29,261		6.90	7.39
Software fusion	-0.00175	£543	11.59	8.40	£28,601	11.59	8.40	£29,144		6.94	7.43
	Inc QALY loss	Inc Costs	Inc LYs*	Inc QALYs*	Inc Costs*	Inc LYs*	Inc QALYs*	Inc Costs*		INHB at £20,000**	INHB at £30,000**
Software fusion vs. cognitive fusion	0.00001	£98	0.04	0.03	-£215	0.04	0.03	-£117	Dominates	0.04	0.03

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 138 Deterministic results for scenario 5 - radical treatment for all identified CPG≥2 and conservative treatment for CPG1 - cost-effectiveness results: combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results			ICER**	NHB at £20,000**	NHB at £30,000**
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*			
Cognitive fusion	-0.00177	£448	11.55	8.36	£28,786	11.55	8.35	£29,234		6.89	7.38
Software fusion	-0.00176	£544	11.63	8.41	£28,390	11.63	8.41	£28,934		6.96	7.44
	Inc QALY loss	Inc Costs	Inc LYs*	Inc QALYs*	Inc Costs*	Inc LYs*	Inc QALYs*	Inc Costs*		INHB at £20,000**	INHB at £30,000**

Software fusion vs. cognitive fusion	0.00002	£95	0.08	0.05	-£396	0.08	0.05	-£300	Dominates	0.07	0.06

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 139 Deterministic results for scenario 6.1 – throughput (150/year) - cost-effectiveness results: targeted biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00176	£495	11.45	8.29	£27,956	11.45	8.29	£28,451		6.87	7.34
Software fusion	-0.00175	£661	11.46	8.30	£27,919	11.46	8.30	£28,580		6.87	7.35
	Inc QALY loss	Inc Costs	Inc LYs*	Inc QALYs*	Inc Costs*	Inc LYs*	Inc QALYs*	Inc Costs*		INHB at £20,000**	INHB at £30,000**
Software fusion vs. cognitive fusion	0.00001	£166	0.02	0.01	-£37	0.02	0.01	£129	£11,425	0.00	0.01

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 140 Deterministic results for scenario 6.1 – throughput (150/year) - cost-effectiveness results: combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00177	£498	11.44	8.28	£27,924	11.44	8.28	£28,422		6.86	7.33
Software fusion	-0.00176	£662	11.49	8.31	£27,870	11.49	8.30	£28,532		6.88	7.35

	Inc QALY loss	Inc Costs	Inc LYs*	Inc QALYs*	Inc Costs*	Inc LYs*	Inc QALYs*	Inc Costs*		INHB at £20,000**	INHB at £30,000**
Software fusion vs. cognitive fusion	0.00002	£164	0.05	0.03	-£54	0.05	0.03	£110	£4,275	0.02	0.02

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 141 Deterministic results for scenario 6.2 – throughput (450/year) - cost-effectiveness results: combined biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00177	£432	11.44	8.28	£27,878	11.44	8.28	£28,309		6.86	7.34
Software fusion	-0.00176	£504	11.49	8.31	£27,831	11.49	8.30	£28,335		6.89	7.36
	Inc QALY loss	Inc Costs	Inc LYs*	Inc QALYs*	Inc Costs*	Inc LYs*	Inc QALYs*	Inc Costs*		INHB at £20,000**	INHB at £30,000**
Software fusion vs. cognitive fusion	0.00002	£73	0.05	0.03	-£47	0.05	0.03	£26	£1,009	0.02	0.02

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit

Table 142 Deterministic results for scenario 6.2 – throughput (450/year) - cost-effectiveness results: targeted biopsy

Strategy	Diagnostic model		Long-term model			Overall results					
	QALY loss	Total Costs	Total LYs*	Total QALYs*	Total Costs*	Total LYs*	Total QALYs*	Total Costs*	ICER**	NHB at £20,000**	NHB at £30,000**
Cognitive fusion	-0.00176	£428	11.45	8.29	£27,907	11.45	8.29	£28,335		6.87	7.34

Software fusion	-0.00175	£503	11.46	8.30	£27,873	11.46	8.30	£28,377		6.88	7.35
	Inc QALY loss	**Inc Costs**	**Inc LYs***	**Inc QALYs***	**Inc Costs***	**Inc LYs***	**Inc QALYs***	**Inc Costs***		**INHB at £20,000****	**INHB at £30,000****
Software fusion vs. cognitive fusion	0.00001	£75	0.02	0.01	-£33	0.02	0.01	£42	£3,689	0.01	0.01

*Discounted at 3.5% per annum; **Per additional QALY; INHB, incremental net health benefit; Inc, incremental; LYs, life years; NHB, net health benefit