National Institute for Health and Care Excellence

Draft

Prostate cancer: diagnosis and management

Evidence reviews for risk stratification of localised prostate cancer

NICE guideline NG131

Evidence reviews underpinning recommendations 1.2.15, 1.2.16, 1.3.7 to 1.3.12, 1.3.21 to 1.3.25 and research recommendation

October 2021

Draft for Consultation

These evidence reviews were developed by Guideline Updates Team

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN:

Contents

Risk stratification of localised prostate cancer

1.1 Review question

In people with localised or locally advanced prostate cancer, which risk stratification models/tools/categorising systems perform better in indicating risk of poor outcomes?

1.1.1 Introduction

The NICE guideline on prostate cancer: diagnosis and management (NICE guideline NG131) was reviewed in 2020 as part of the NICE's surveillance programme. New evidence was identified which suggested that the 3 criteria model for risk stratification used in recommendation 1.2.16 of NG131 could be out of date.

Currently, recommendation 1.2.16 provides a table of risk stratification for people with localised prostate cancer. This model stratifies people into low, intermediate and high risk based on 3 criteria: prostate-specific antigen, Gleason score and clinical stage. The subsequent treatment recommendations based on this risk stratification, particularly around active surveillance, were based on longitudinal studies and committee consensus. A new model for risk stratification (Cambridge Prognostic Group [CPG]) stratifies people into low risk (CPG1), favourable intermediate risk (CGP2), unfavourable intermediate risk (CPG3), high risk (CPG4), and very high risk (CPG5).

The new evidence identified by the NICE's surveillance programme indicated that active surveillance may not be appropriate in patients with unfavourable intermediate prostate cancer, and that there may be over treatment of favourable intermediate risk and lower risk patients. Recommendation 1.2.16 is based on the 3-tier risk stratification and it does not differentiate between favourable intermediate risk (CPG2) and unfavourable intermediate risk (CPG3), unlike the CPG criteria. Furthermore, the National Prostate Cancer Audit (NPCA) is now moving to use the 5-tier CPG criteria also means that NG131 will be out of step with key UK auditing and system improvement measures.

It was concluded that this new evidence is a sufficient basis for an expert committee to consider the impact on risk stratification (recommendation 1.2.16) and the subsequent treatment recommendations.

The aim of this review is to assess which risk stratification models/tools/categorising systems perform better in indicating risk of poor outcomes in people with localised or locally advanced prostate cancer. This review identified retrospective cohort studies that fulfilled the conditions specified in Table 1. See Appendix A for full details of the review protocol.

1.1.2 Summary of the protocol

Table 1: PICO table for risk stratification of localised prostate cancer

Population	**Inclusion** People newly diagnosed with localised/locally advanced prostate cancer **Exclusion** People diagnosed with metastatic cancer (including oligometastatic cancer) as documented by M stage disease and/or positive bone or CT scan
Predictor	• 5-tier prostate cancer risk stratification tools (for example Cambridge Prognostic Group (CPG)) • 3 tier prostate cancer risk stratification tools (for example NICE's tool)

Comparator	• 5-tier prostate cancer risk stratification tools (for example Cambridge Prognostic Group (CPG)) • 3 tier prostate cancer risk stratification tools (for example NICE's tool)
Outcome (s)	**Clinical endpoints** • Progression to metastatic prostate cancer • Progression free survival (including radiological and biochemical progression free survival) • Metastases free survival • Prostate cancer specific mortality • Health related quality of life For each outcome, metrics measures will be reported where available, for example: • Odds ratios/hazard ratios • Model fit statistics (for example R2, Brier score) • Discrimination (for example C statistic, area under ROC curve) • Calibration (for example calibration slope)

The Memorial Sloan Kettering Cancer Centre (MSKCC) and the Cancer of the Prostate Risk Assessment (CAPRA) were in the original protocol as examples (see appendix A) but both were subsequently removed because neither of them are 3 or 5 tier models.

1.1.3 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in Appendix A and the methods section in Appendix K.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

The committee first considered evidence on risk stratification tools outlined in this evidence review document. After making recommendations on the risk stratification tool that should be used, the committee considered the impact of this recommendation on other recommendations in the NICE prostate cancer guideline. The committee amended recommendations that referred to the previous classification scheme, taking into account the original evidence that the recommendations were based on and their knowledge and experience.

The 2019 evidence review comparing active surveillance to radical treatment for localised prostate cancer was used to inform recommendations on treatment options for localised prostate cancer. The 2014 evidence review on hormone therapy was used to inform recommendations on hormone treatments in combination with radical radiotherapy. The 2008 review on bone scans was used to inform recommendations on bone scans in people with newly diagnosed prostate cancer. The 2019 review on radiotherapy was used to inform recommendations on brachytherapy.

1.1.4 Prognostic evidence

1.1.4.1 Included studies

A systematic search was carried out to identify prognostic observational studies and systematic reviews of these studies, which found 8,933 references (see Appendix B for the literature search strategy). Evidence from studies referenced in identified systematic reviews were also reviewed (2 references from a systematic review were not found by the search). In total, 8,935 references were identified for screening at title and abstract level using priority

screening. From the first 4,720 references screened, 4,694 were excluded based on their titles and abstracts and 26 references were ordered for screening based on their full texts. Based on the rules for using priority screening software (see Appendix K), the screening was terminated at this point, and the remaining 4,215 references were not screened on title and abstract.

Of the 26 references screened as full texts, 5 references (all retrospective cohort studies) were included based on their relevance to the review protocol (Appendix A). The clinical evidence study selection is presented as a diagram in Appendix C.

See section s 1.2.15, 1.2.16, 1.3.7 to 1.3.12, 1.3.21 to 1.3.25 and the research recommendation for a list of included references.

1.1.4.2 Excluded studies

See Appendix J for a list of excluded studies with reasons for exclusion.

1.1.5 Summary of studies included in the prognostic evidence

Table 2: Summary of studies on risk stratification models in people with localised or locally advanced prostate cancer (see Appendix L for details on each model)

Study	Population	Risk stratification models	Outcomes
Abdel-Rahman 2018	Men with N0/M0 disease according to the TNM sixth system Validation cohort from the US (n=30,445)	• D'Amico • Modified risk stratification model (incorporation of percent of positive cores into D'Amico)	• C-statistic for prostate cancer specific mortality
Gnanapragasam 2016	Cases with all components of diagnostic stage, primary and secondary grade, and presenting PSA as well as data on follow-up and survival Validation cohort from Northern Ireland (n=1,706)	• NICE • CPG	• C-statistic for prostate cancer specific mortality
Gnanapragasam 2018	Men with no evidence of metastatic disease (Mx or M0) and with PSA <100 ng/ml Sweden cohort (n=72,337) Singapore cohort (n=2,550)	• NICE • CPG	• Hazard ratios for prostate cancer specific mortality
Lee 2021	Men aged 35 to 95 years diagnosed with histologically confirmed non-metastatic prostate cancer Participants from the US (n=171,942)	• CPG • EAU • GUROC • NICE	• C-statistic for prostate cancer specific mortality • Brier score
Zelic 2020	Men diagnosed with	• AUAi	• Hazard ratios for

Study	Population	Risk stratification models	Outcomes
	non-metastatic (not M1 or N1) prostate cancer Participants from Sweden (n=139,515)	• CPG • EAU • GUROC • NICE	prostate cancer specific mortality • C-statistic for prostate cancer specific mortality

(a) American Urological Association with favourable and non-favourable intermediate groups (AUA-i), Cambridge Prognostic Groups (CPG), European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence)

See Appendix D for full evidence tables.

1.1.6 Summary of the prognostic evidence

Table 3: 3 tier prostate cancer risk stratification models for prediction of prostate cancer specific mortality

Risk stratification tier	Reference	No. of participants	Hazard ratio (95% CI)	Quality
NICE risk stratification model				
Intermediate risk	Low risk	139,515	2.94 (2.51, 3.44)	Moderate
NICE risk stratification model				
High risk	Low risk	139,515	14.16 (12.42, 16.14)	Moderate
D'Amico risk stratification model				
Intermediate risk	Low risk	139,515	2.88 (2.45, 3.38)	Moderate
D'Amico risk stratification model				
High risk	Low risk	139,515	13.69 (12.00, 15.62)	Moderate
EAU risk stratification model				
Intermediate risk	Low risk	139,515	2.94 (2.51, 3.44)	Moderate
EAU risk stratification model				
High risk	Low risk	139,515	14.16 (12.42, 16.14)	Moderate
GUROC risk stratification model				
Intermediate risk	Low risk	139,515	3.22 (2.77, 3.76)	Moderate
GUROC risk stratification model				
High risk	Low risk	139,515	16.08 (14.10, 18.35)	Moderate

European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence)

Table 4: 5 tier prostate cancer risk stratification models for prediction of prostate cancer specific mortality

Risk stratification tier	Reference	No. of participants	Hazard ratio (95% CI)	Quality
CPG risk stratification model				
CGP2	CGP1	178,969	2.32	Moderate

Risk stratification tier	Reference	No. of participants	Hazard ratio (95% CI)	Quality
			(2.11, 2.55)	
CPG risk stratification model				
CGP3	CGP1	173,019	4.63 (4.17, 5.13)	Moderate
CPG risk stratification model				
CGP4	CGP1	179,488	7.79 (7.20, 8.43)	Moderate
CPG risk stratification model				
CGP5	CGP1	177,036	22.72 (18.83, 27.42)	Very low
CPG risk stratification model				
CGP2	CGP1	25,303	2.30 (2.04, 2.59)	Moderate
CPG risk stratification model				
CGP3	CGP2	14,796	2.11 (1.89, 2.36)	Moderate
CPG risk stratification model				
CGP4	CGP3	7,354	1.56 (1.42, 1.72)	Moderate
CPG risk stratification model				
CGP5	CGP4	13,506	2.72 (2.58, 2.88)	Moderate
AUA-i risk stratification model				
Low risk	Very low risk	139,515	1.11 (0.83, 1.49)	Low
AUA-i risk stratification model				
Favourable intermediate risk	Very low risk	139,515	2.54 (2.00, 3.23)	Moderate
AUA-i risk stratification model				
Unfavourable intermediate risk	Very low risk	139,515	5.15 (4.05, 6.55)	Moderate
AUA-i risk stratification model				
High risk	Very low risk	139,515	17.64 (14.12, 22.05)	Moderate

American Urological Association with favourable and non-favourable intermediate groups (AUA-i), Cambridge Prognostic Groups (CPG)

Table 5: Validity of 3 tier prostate cancer risk stratification models for prediction of prostate cancer specific mortality – Discrimination (c-statistic)

Risk stratification model	Follow-up	Study (s)	No. of participants	C-statistic (95% CI)	Quality
NICE risk stratification model	Median 5.9 years	Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort	248,535	0.73 (0.68, 0.77)	Very low

Risk stratification model	Follow-up	Study (s)	No. of participants	C-statistic (95% CI)	Quality
		Lee 2021			
NICE risk stratification model – sensitivity analysis without studies at high risk of bias	Median 4.8 years	Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort	76,593	0.73 (0.66, 0.80)	Very low
NICE risk stratification model	10 years	Zelic 2020	139,515	0.73	Moderate
D'Amico risk stratification model	10 years	Zelic 2020	139,515	0.73 (0.72, 0.73)	Moderate
D'Amico risk stratification model	Median 2.25 years	Abdel-Rahman 2018	30,445	0.78 (0.75, 0.81)	Very low
EAU risk stratification model	10 years	Lee 2021	171,942	0.71 (070, 0.72)	Low
GUROC risk stratification model	10 years	Lee 2021	171,942	0.75 (0.73, 0.76)	Low

European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence)

Table 6: Validity of 5 tier prostate cancer risk stratification models for prediction of prostate cancer specific mortality – Discrimination (c-statistic)

Risk stratification model	Follow-up	Study (s)	No. of participants	C-statistic (95% CI)	Quality
CPG risk stratification model	Median 7 years	Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort Lee 2021 Zelic 2020	388,050	0.79 (0.77, 0.81)	Very low
CPG risk stratification model – sensitivity analysis without studies at high risk of bias	Median 5.9 years	Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort Zelic 2020	216,108	0.79 (0.77, 0.82)	Very low
D'Amico risk stratification model*	Median 2.25 years	Abdel-Rahman 2018	30,445	0.81 (0.78, 0.84)	Very low

** Incorporation of percent of positive cores, Cambridge Prognostic Groups (CPG)*

Table 7: Validity of 3 tier prostate cancer risk stratification models for prediction of prostate cancer specific mortality – Calibration (Brier score)

Risk stratification model	Follow-up	Study (s)	Sample size	Quality	Brier score (95% CI)[a,b]
NICE	10 years	Lee 2021	171,942	Low	0.039 (0.037, 0.041)
EAU	10 years	Lee 2021	171,942	Low	0.039 (0.037, 0.041)
GUROC	10 years	Lee 2021	171,942	Low	0.039 (0.037, 0.041)

(a) Lower numbers (closer to zero) reflect better calibration (and therefore predictive accuracy)
(b) The median difference between observed vs predicted prostate cancer specific mortality
European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence)

Table 8: Validity of 5 tier prostate cancer risk stratification models for prediction of prostate cancer specific mortality – Calibration (Brier score)

Risk stratification model	Follow-up	Study (s)	Sample size	Quality	Brier score (95% CI)[a,b]
CPG	10 years	Lee 2021	171,942	Low	0.037 (0.035, 0.039)

(a) Lower numbers (closer to zero) reflect better calibration (and therefore predictive accuracy)
(b) The median difference between observed vs predicted prostate cancer specific mortality
Cambridge Prognostic Groups (CPG)

See Appendix F for full GRADE tables.

1.1.7 Economic evidence

1.1.7.1 Included studies

A systematic review was conducted to identify economic evaluations for this review question. The search returned 151 records which were sifted against the review protocol. All of these studies were excluded based on title and abstract.

1.1.7.2 Excluded studies

All studies were excluded at title and abstract screening.

1.1.8 Summary of included economic evidence

Any change in the risk classification tool recommended (i.e. from the NICE 3-tier tool to the CPG 5 tier tool) would require updates to any existing recommendations that are based on patients being assigned one of those risk categories. The evidence used previously to underpin those recommendations was revisited with the committee to confirm that all recommendations would still hold when the current risk categories were switched to the CPG risk levels.

Recommendations on treatment options for localised prostate cancer (2019)

Evidence from studies by Koerber (2014), Ramsay (2015), and Lyth (2012) was used in the 2019 prostate cancer guideline update to underpin recommendations on treatment for localised prostate cancer.

The studies by Koerber and Ramsay were economic evaluations in low-risk prostate cancer populations corresponding to those with PSA ≤ 10, Gleason score ≤ 6 and T stage ≤ T2a. Koerber et al. found that in this low-risk group radical prostatectomy was dominated by active surveillance. Ramsay et al. conducted a sensitivity analysis where radical treatments were compared against active surveillance which was found to be dominant over radical treatments. A limitation to the analysis conducted by Ramsay et al. was that data and assumptions were used in the absence of direct data to construct the active surveillance comparator, so the result should be treated with caution.

Lyth et al. compared watchful waiting to radical prostatectomy in various age and risk groups; PSA ≤ 10 and Gleason score ≤ 6, PSA 11-20 and Gleason score 7, or PSA > 20 and Gleason score ≥ 8, in ages 65, 70, and 75. In these analyses radical prostatectomy was found to be more cost-effective than watchful waiting with an ICER below the 200,000 SEK (£17,000) per QALY threshold in all groups other than the low-risk 75 years group.

Recommendations on hormone therapy (2014)

In the absence of economic evidence on hormone therapy when these recommendations were developed for the 2014 prostate cancer guideline, the committee used clinical experience and consensus to estimate the resource use associated with the recommendations.

Recommendations on brachytherapy (2019)

Three economic evaluations on brachytherapy were identified for the 2019 prostate cancer guideline update; Ollendorf et al. 2008, Ramsay et al. 2015, and Sanyal et al. 2016. All three studies were judged to have potentially serious or very serious limitations.

Ollendorf et al. compared brachytherapy with both proton beam therapy (PBT) and intensity modulated radiotherapy (IMRT) in low-risk prostate cancer, defined as stage T1-T2a lesions, Gleason score 2-6, and PSA ≤ 10. Ollendorf found that brachytherapy was dominant over PBT and IMRT in this low-risk population.

Ramsay et al. compared brachytherapy with IMRT in a mixed risk population and found that brachytherapy was not cost-effective in this comparison, with an ICER of £84,883.

Sanyal et al. conducted two analyses; brachytherapy plus IMRT compared with IMRT alone in intermediate-risk prostate cancer, and brachytherapy compared with IMRT in low-risk prostate cancer. Low-risk disease corresponds to stage T1-T2a, Grade group 1, and PSA ≤ 10, and intermediate-risk disease corresponds to stage T2b-T2c, Grade group 2-3, PSA 10-20. Sanyal found that in low-risk prostate cancer IMRT is dominated by brachytherapy, and in intermediate-risk prostate cancer the combination of brachytherapy and IMRT is dominated by IMRT alone.

Recommendations on bone scans

The committee did not rate this topic as a priority for health economics, therefore economic evidence on bone scans was not reviewed when this recommendation was made during development of the 2008 prostate cancer guideline.

1.1.9 Economic model

No original economic modelling was completed for this review question.

1.1.10 Evidence statements

No existing economic studies or *de novo* economic modelling was included in this review question.

1.1.11 The committee's discussion and interpretation of the evidence

1.1.11.1. The outcomes that matter most

The committee agreed that prostate cancer specific mortality was an important outcome in people newly diagnosed with localised or locally advanced prostate cancer. Other outcomes were also considered to be important (progression to metastatic prostate cancer, progression free survival, metastases free survival, and health related quality of life) but no evidence was found reporting on any of these outcomes. However, prostate cancer specific mortality is likely to be highly correlated with these other measures.

This review used 3 groups of measures to assess prognostic accuracy of the 3 tier and 5 tier risk prediction models. These measures were hazard ratios, c-statistics and Brier scores. The 3 measures were used to assess the performance of the risk stratification models in predicting prostate cancer mortality in people with localised or locally advanced prostate cancer.

1.1.11.2 The quality of the evidence

Overall, the quality of the evidence varied from moderate to very low, with the main reasons for downgrading being due to the lack of information on whether the risk stratification models were calculated without the knowledge of the outcome (prostate cancer specific mortality) and the lack of clarify on whether competing risk analyses were used in predicting prostate cancer specific mortality. Imprecision and heterogeneity were also reasons for downgrading the evidence. In some of the hazard ratios and c-statistics results, imprecision was considered to be serious because the 95% confidence intervals crossed the line of no effect (for hazard ratios) or because the 95% confidence intervals crossed 2 categories of test classification accuracy (c-statistic; see Table 10 for details on classification accuracy). Heterogeneity was considered to be very serious when comparing 2 of the tiers of the Cambridge Prognostic Groups (CPG) risk stratification model with a I^2 >66.7% (CPG5 compared to CPG1). Meta-analyses combining either 3 tier models or 5 tier risk stratification models also showed very serious heterogeneity with a I^2 >66.7%.

The committee acknowledge that evidence on c-statistics for the NICE 3 tier and CPG 5 model had limitations in terms of imprecision and heterogeneity. However, the high heterogeneity was largely due to the narrow confidence intervals for some individual studies and the differences in c-statistics between studies were small. Evidence on c-statistics was also downgraded for imprecision, although the confidence intervals for the overall result were narrow, as they crossed the pre-specified categories for c—statistic performance. However, this imprecision was not a major concern for decision making.

No evidence was found reporting on the rest of the outcomes listed in the protocol (progression to metastatic prostate cancer, progression free survival, metastases free survival, and health related quality of life).

1.1.11.3 Discussions about risk stratification models for people with localised or locally advanced prostate cancer

New evidence showed that 5-tier risk prediction models discriminate better when predicting prostate cancer specific mortality compared to 3-tier models. This was shown by higher c-statistics (c-statistic from 0.78 to 0.81 for 5 tier models and from 0.66 to 0.78 for 3 tier models). The 5 tier risk stratification models also showed marginally better calibration with a

lower Brier score of 0.037 for the 5-tier CPG model compared to 0.039 for 3 tier models that were assessed. Despite an overlap in confidence intervals, the committee noted that the 5 tier CPG model made sense in terms of their clinical experience of disease progression for people with different grade groups. The previous NICE 3-tier model put all people with a Gleason score of 7 into the intermediate risk category, whereas the CPG model takes into account the Grade group, which distinguishes between a Gleason score of 3+4=7 and a score of 4+3=7, which is known to have a different prognosis.

The risk of prostate cancer specific mortality was significantly higher in higher risk groups for both 3 tier and 5 tier models (highest hazard ratio was 16.08 for 3 tier models and 22.72 for 5 tier models). There was also evidence of significant and steadily increase in the risk of prostate cancer specific mortality when adjacent tiers of the CPG model were compared to each other with hazard ratios from 2.3 to 2.7 (CPG2 compared to CPG1; CGP3 compared to CGP2; CPG4 compared to CPG4; and CPG5 compared to CPG4).

The committee discussed the different 5 tier models and noted that all of the 5-tier models apart from the CPG model included the assessment of percentage of positive cores and cores with percentage of cancer. The committee agreed that there are limitations when assessing the percentage of cores involved when using MRI-guided biopsy, which is the current recommended way of diagnosing and assessing prostate cancer. Therefore, the 5 tier CPG model was recommended because this model does not assess percentage of positive cores and cores with percentage of cancer as part of the calculation of the tiers. Additionally, the same information is needed for calculating either the 3 tier NICE risk stratification model and the 5 tier CPG model which means that there would not be a resource impact in clinical practice at calculating the 5 tier CPG model compared to the 3 tier NICE model. The committee noted that the evidence for the CPG 5 tier model was from a UK study and therefore tested on a UK population.

The committee agreed that 5 tier models break down intermediate and high risk groups into subgroups which provides more clarity regarding of the treatment pathway for each of the subgroups. This in turn might reduce under and over treatment in people who are at either end of the tiers. The committee noted that in their clinical experience, using the CPG model might prevent over treatment in people with lower risk of prostate cancer specific mortality. This might mean that fewer people would have unnecessary radical treatment.

The committee discussed about the terminology to use when referring to the pathology information needed to calculate the tiers of the CPG model. They agreed that it was important to keep both the Gleason score and the grade group (pathological classification of prostate cancer from the International Society of Urological Pathology [ISUP]). The committee also highlighted that most clinicians talk about ‘grade group’ rather than ISUP. Therefore, the committee agreed to use the term ‘grade group’ in the recommendation.

1.1.11.4 Cost effectiveness and resource use

No economic evidence was identified comparing 3-tier and 5-tier risk tools however, the committee was confident that recommending the 5-tier CPG risk stratification model would not have a significant resource impact as the same information is used to calculate both the CPG model and the previously recommended 3-tier model.

The committee did not expect that the changes to the existing recommendations resulting from use of the 5-tier CPG risk stratification tool would have any significant resource impact.

The economic evidence used in previous versions of the guideline to make recommendations based on patient risk was still agreed to be relevant and the committee considered this evidence when these recommendations were updated to use the 5-tier CPG risk stratification tool.

The committee did not think that the changes to the risk stratification tool would lead to a change in practice regarding the recommendations on hormone therapy so the only changes made to those recommendations were amendments to the terminology used to define the risk groups.

Recommendations where low-risk was replaced with CPG 1 are likely to apply to a broader population, however, the committee agreed that the associated resource impact of this change would be minimal because although the recommendation places more emphasis on active surveillance, the other treatment options are still available to those people. Offering active surveillance to people with CPG 1 prostate cancer, and only considering radical treatment where active surveillance is unsuitable or unacceptable may plausibly be cost saving but it is not possible to quantify this as it depends on the individual person and their treatment pathway and disease progression.

Recommendations where intermediate-risk was replaced with CPG 2 and CPG 3 are likely to apply to a smaller group of people, and the committee agreed that the changes were unlikely to result in an increased use of resources.

Recommendations that were previously for high-risk prostate cancer were changed to be for CPG 4 and CPG 5, and since these groups are equivalent there would be no resource impact in using the CPG risk tool.

1.1.11.5 Other factors the committee took into account

The committee considered the impact of recommending a 5-tier risk stratification model on existing treatment recommendations that refer to the 3-tier model elsewhere in the guideline. The recommendations affected by this change were on:

- Radical radiotherapy, radical prostatectomy, and active surveillance
- Isotope bone scans
- Hormone therapy
- Brachytherapy

Radical radiotherapy, radical prostatectomy, and active surveillance

When considering the 2019 recommendation to offer a choice of radical treatment or active surveillance to those in a low-risk tier, the committee agreed that ‘low risk’ could be mapped to CPG 1 in the 5-tier model but the population in CPG1 was slightly broader, encompassing some intermediate risk people. When evaluating the evidence used to underpin the 2019 recommendation, the committee noted that the Protec T trial (which was based in a UK in the UK and whose population was most similar to the CPG 1 risk category) showed no benefit in choosing radical treatment over active-surveillance in relation to mortality outcomes and that adverse events in treatment groups were much higher. Given this interpretation of the evidence, the committee felt strongly that active surveillance should be offered as the preferred option to patients in this group, but that treatment should be considered for patients in whom active surveillance was unacceptable. This change in emphasis also matched their experience of what was happening in UK practice and addressed wider concerns about overtreatment in people with low-risk of disease progression. The committee also felt that patients would feel reassured in choosing active surveillance in the knowledge that this was the preferred option recommended by NICE. From a patient perspective, committee noted that in their experience many people regretted having radical treatment and that presenting active surveillance and radical treatment as equal options to people in the CPG 1 group who have low risk of disease progression is misleading.

When considering the 2019 recommendation on offering treatment and considering active surveillance for those in the intermediate risk group, the committee agreed that people in the CPG2 group should be offered all three options. The committee discussed offering all three options to people in the CPG3 group as this also mapped to an intermediate risk in the 3-tier

model. The committee highlighted however that the CPG3 group contained people with Gleason pattern 4+3 (grade group 3) in whom active surveillance would not be the preferred clinical option given its association with poorer outcomes. Balanced against this, the committee also highlighted that the evidence used to underpin the 2019 recommendation did include some people in the higher-intermediate/CPG3 risk group on active surveillance and to remove considering active surveillance for these people was not supported by the evidence that had been reviewed. The committee agreed that two recommendations should be drafted to place the emphasis on offering treatment as the preferred option to those in the CPG3 group but to consider active surveillance for those in whom treatment was unacceptable.

When considering the 2019 recommendations in high-risk groups, the committee agreed active surveillance should not be considered as an option, and that CPG4 and CPG5 were equivalent to the 'high risk' group in the previous recommendations.

Hormone therapy

The committee agreed that the 2014 recommendation to offer hormone therapy in combination with radical radiotherapy to intermediate and high-risk groups could be amended to CPG 2-5 from the 5-tier model. The committee agreed that the recommendation to consider continuing hormone therapy for up to 3 years in high-risk groups could be amended to CPG 4-5. The committee agreed that these CPG groups were broadly equivalent to intermediate and high risk and the populations receiving radical radiotherapy, and that this amendment would not constitute a change in current practice.

Brachytherapy

The committee agreed that the 2019 recommendation to consider brachytherapy in combination with radiotherapy in people with intermediate and high risk localised prostate cancer could be amended to CPG2-5 as these were broadly equivalent to intermediate and high-risk groups. Similarly, the 2008 recommendation not to offer brachytherapy alone to high-risk groups could be amended to CPG4-5 groups.

Bone Scan

The committee agreed that the 2019 recommendation that bone scans should not be used for people with low-risk prostate cancer could be amended for to the CPG1 population. The committee were aware that this population is broader than the low-risk population referred to in the previous guideline but agreed that it was in line with current practice not to offer bone scans to this group.

The committee also highlighted the lack of evidence for the CPG 2 and 3 groups on when to offer staging investigations more generally and the potential resource impact of these investigations and made a research recommendation for these groups.

1.1.12 Recommendations supported by this evidence review

This evidence review supports recommendations 1.2.15, 1.2.16, 1.3.7 to 1.3.12, 1.3.21 to 1.3.25 and the research recommendation on staging investigations for CPG 2 and 3 prostate cancer.

1.1.13.1 Prognostic evidence

Abdel-Rahman, Omar (2018) Dissecting the heterogeneity of localized prostate cancer risk groups through integration of percent of positive cores. Future oncology (London, England) 14(15): 1469-1476

Gnanapragasam, V J, Bratt, O, Muir, K et al. (2018) The Cambridge Prognostic Groups for improved prediction of disease mortality at diagnosis in primary non-metastatic prostate cancer: a validation study. BMC medicine 16(1): 31

Gnanapragasam, Vincent J, Lophatananon, Artitaya, Wright, Karen A et al. (2016) Improving Clinical Risk Stratification at Diagnosis in Primary Prostate Cancer: A Prognostic Modelling Study. PLoS medicine 13(8): e1002063

Lee, Changhee, Light, Alexander, Alaa, Ahmed et al. (2021) Application of a novel machine learning framework for predicting non-metastatic prostate cancer-specific mortality in men using the Surveillance, Epidemiology, and End Results (SEER) database. The Lancet. Digital health 3(3): e158-e165

Zelic, Renata, Garmo, Hans, Zugna, Daniela et al. (2020) Predicting Prostate Cancer Death with Different Pretreatment Risk Stratification Tools: A Head-to-head Comparison in a Nationwide Cohort Study. European urology 77(2): 180-188

1.1.13.2 Economic

Koerber, F., Waidelich, R., Stollenwerk, B. et al (2014) The cost-utility of open prostatectomy compared with active surveillance in early localised prostate cancer. BMC Health Serv Res 14, 163

Lyth J, Andersson SO, Andren O, Johansson JE, Carlsson P, Shahsavar N. (2012) A decision support model for cost-effectiveness of radical prostatectomy in localized prostate cancer. Scandinavian Journal of Urology and Nephrology 46(1): 19-25

Ollendorf, D., Hayes, J., McMahon, P., Pearson, S., Kuba, M., & Tramontano, A. (2008) Institute for Clinical and Economic Review Final Appraisal Document: Brachytherapy and Proton Beam Therapy for Treatment of Clinically Localized, Low-Risk Prostate Cancer.

Ramsay CR, Adewuyi T, Gray J, Hislop J, Shirley MDF, Jayakody S, et al. (2015) Ablative therapy for people with localised prostate cancer: a systematic review and economic evaluation. Health Technol Assess 19(49)

Sanyal C, Aprikian AG, Cury FL, Chevalier S, Dragomir A. (2016) Management of localized and advanced prostate cancer in Canada: A lifetime cost and quality-adjusted life-year analysis. Cancer 122(7):1085-96

Appendices

Appendix A – Review protocols

Review protocol for risk stratification of localised prostate cancer

ID	Field	Content
0.	PROSPERO registration number	CRD42021270616
1.	Review title	Staging – risk stratification tools for localised prostate cancer.
2.	Review question	In people with localised or locally advanced prostate cancer, which risk stratification models/tools/categorising systems perform better in indicating risk of poor outcomes?
3.	Objective	To determine if the 5 tier-risk stratification models/tools/categorising systems perform better in indicating risk of poor outcomes compared to the currently recommended 3-tier risk stratification model for people with localised prostate cancer.
4.	Searches	The following databases will be searched: • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE

		Searches will be restricted by: • 2007 • English language Other searches: • Reference searching • Citation searching • Inclusion lists of systematic reviews The full search strategies for MEDLINE database will be published in the final review.
5.	Condition or domain being studied	Risk stratification tools for localised/locally advanced prostate cancer.
6.	Population	Inclusion: People newly diagnosed with localised/locally advanced prostate cancer. Exclusion: • People diagnosed with metastatic cancer (including oligometastatic cancer) as documented by M stage disease and/or positive bone or CT scan.

7.	Intervention/Exposure/Test	• 5-tier prostate cancer risk stratification tools (for example Cambridge Prognostic Group (CPG), Memorial Sloan Kettering Cancer Centre (MSKCC), Cancer of the Prostate Risk Assessment (CAPRA) • 3 tier prostate cancer risk stratification tools (for example NICE's tool)
8.	Comparator/Reference standard/Confounding factors	• 5-tier prostate cancer risk stratification tools (for example Cambridge Prognostic Group (CPG), Memorial Sloan Kettering Cancer Centre (MSKCC), Cancer of the Prostate Risk Assessment (CAPRA) • 3 tier prostate cancer risk stratification tools (for example NICE's tool)
9.	Types of study to be included	• Prospective cohort studies • Retrospective cohort studies • Model validation studies • Model impact studies • Systematic reviews of these studies
10.	Other exclusion criteria	• All other study types. • Model development studies that do not report model validation data.
11.	Context	NICE guideline NG131 recommendations on risk stratification of prostate cancer will be updated by this review question.
12.	Primary outcomes (critical outcomes)	Clinical endpoints • Progression to metastatic prostate cancer.

		• Progression free survival (including radiological and biochemical progression free survival). • Metastases free survival. • Prostate cancer specific mortality. • Health related quality of life For each outcome, metrics measures will be reported where available, for example: • Odds ratios/hazard ratios • Model fit statistics (for example R^2, Brier score) • Discrimination (for example C statistic, area under ROC curve). • Calibration (for example calibration slope)
13.	Secondary outcomes (important outcomes)	Not applicable
14.	Data extraction (selection and coding)	All references identified by the searches and from other sources will be uploaded into EPPI reviewer and de-duplicated. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. A standardised form will be used to extract data from studies (see Developing NICE guidelines: the manual section 6.4). Study investigators may be contacted for missing data where time and resources

		allow. This review will make use of the priority screening functionality within the EPPI-reviewer software.
15.	Risk of bias (quality) assessment	Risk of bias will be assessed using the PROBAST checklist as described in Developing NICE guidelines: the manual.
16.	Strategy for data synthesis	**Approach to meta-analysis** Where appropriate, hazard ratios will be pooled using the generic inverse-variance method. Adjusted odds ratios, hazard ratios and risk ratios from multivariate models will only be pooled if the same set of factors are used across multiple studies and if the same thresholds to measure factors were used across studies. Meta-analysis of c statistics will be considered when the same prognostic models have been evaluated across multiple studies. Meta-analyses of c statistics will be carried out using the metamisc package in R v3.4.0, which confines the analysis results to between 0 and 1 matching the limited range of values that c-statistics can take. Random effects meta-analysis will be used when the I^2 is 50% or greater. **Approach to GRADE** A modified approach will be applied using the GRADE framework. Evidence from cohort will initially be rated as high-quality, and then assessed according to the same criteria as described in the standard GRADE criteria (risk of

		bias, inconsistency, imprecision and indirectness).
17.	Analysis of sub-groups	None
18.	Type and method of review	☐ Intervention ☐ Diagnostic ☒ Prognostic ☐ Qualitative ☐ Epidemiologic ☐ Service Delivery ☐ Other (please specify)
19.	Language	English
20.	Country	England
21.	Anticipated or actual start date	12/05/2021
22.	Anticipated completion date	To be determined

23.	Stage of review at time of this submission	**Review stage**	**Started**	**Completed**
		Preliminary searches	☐ ☐	☐ ☐
		Piloting of the study selection process	☐ ☐	☐ ☐
		Formal screening of search results against eligibility criteria	☐ ☐	☐ ☐
		Data extraction	☐ ☐	☐ ☐
		Risk of bias (quality) assessment	☐ ☐	☐ ☐
		Data analysis	☐ ☐	☐ ☐
24.	Named contact	**5a. Named contact** NICE Guideline Updates Team **5b Named contact e-mail**		

		[Guideline email]@nice.org.uk [Developer to check with Guideline Coordinator for email address] **5e Organisational affiliation of the review** National Institute for Health and Care Excellence (NICE) and NICE Guideline Updates Team
25.	Review team members	[Give the title, first name, last name and the organisational affiliations of each member of the review team. Affiliation refers to groups or organisations to which review team members belong.] From the [Insert Development centre]: • [Tech lead] • [Tech analyst] • [Health economist] • [Information specialist] • [Others]
26.	Funding sources/sponsor	This systematic review is being completed by the Guideline Updates team which receives funding from NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential

		conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: [NICE guideline webpage].
29.	Other registration details	This is a new review that will update the risk stratification of localised prostate cancer section in the NICE guideline NG131 Prostate Cancer: diagnosis and management (2019.
30.	Reference/URL for published protocol	
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.
32.	Keywords	prostate cancer, non-metastatic cancer, localised prostate cancer, risk stratification models
33.	Details of existing review of same topic by same authors	This is a new review question that will update prostate cancer: diagnosis and management

		(2019) NICE guideline NG131.
34.	Current review status	☒ Ongoing ☐ Completed but not published ☐ Completed and published ☐ Completed, published and being updated ☐ Discontinued
35..	Additional information	N/A
36.	Details of final publication	www.nice.org.uk

1

2

3

Appendix B – Literature search strategies

Search design and peer review

A NICE information specialist conducted the literature searches for the evidence review. The searches were originally run in July 2021. This search report is compliant with the requirements of PRISMA-S.

The MEDLINE strategy below was quality assured (QA) by a trained NICE information specialist. All translated search strategies were peer reviewed to ensure their accuracy. Both procedures were adapted from the 2016 PRESS Checklist.

The principal search strategy was developed in MEDLINE (Ovid interface) and adapted, as appropriate, for use in the other sources listed in the protocol, taking into account their size, search functionality and subject coverage.

Review Management

The search results were managed in EPPI-Reviewer v5. Duplicates were removed in EPPI-R5 using a two-step process. First, automated deduplication is performed using a high-value algorithm. Second, manual deduplication is used to assess 'low-probability' matches. All decisions made for the review can be accessed via the deduplication history.

Prior work

The terms for 'prostate cancer' are based on those used for the previous NICE guideline, NG131 Prostate cancer: diagnosis and management (2019). However, amendments were made to the search strategy as appropriate for this specific evidence review topic.

Limits and restrictions

English language limits were applied in adherence to standard NICE practice and the review protocol.

Limits to exclude, (comment or letter or editorial or historical articles or conference abstract or conference paper or "conference review" or letter or case report) were applied in adherence to standard NICE practice and the review protocol.

The search was limited from January 2007 to July 2021 as defined in the review protocol.

The limit to remove animal studies in the searches was the standard NICE practice, which has been adapted from: Dickersin, K., Scherer, R., & Lefebvre, C. (1994).

Systematic Reviews: Identifying relevant studies for systematic reviews. *BMJ*, 309(6964), 1286.

Search filters

Prognosis

The following search filter was applied to the clinical searches in MEDLINE and Embase to identify prognostic studies*:* McMaster Prognosis – (maximizes sensitivity)

The following terms were also applied from the clinical prediction filter, scor:.tw or observ:mp: McMaster Clinical Prediction Guides – (maximizes sensitivity)

Cost effectiveness searches

The NICE cost utility filter was applied to the search strategies in MEDLINE and Embase to identify cost-utility studies. *(Hubbard W, et al. Development of a validated search filer to identify cost utility studies for NICE economic evidence reviews. NICE Information Services.)*

Clinical searches

Databases	Date searched	Version/files	No. of results downloaded
Cochrane Central Register of Controlled Trials (CENTRAL) via Wiley	27/07/2021	Issue 7 of 12, July 2021	1346
Cochrane Database of Systematic Reviews (CDSR) via Wiley	27/07/2021	Issue 7 of 12, July 2021	0
Database of Abstracts of Reviews of Effect (DARE) via CRD	27/07/2021	n/a	90
Embase (Ovid)	27/07/2021	1974 to 2021 July 26	7604
Health Technology Assessment (HTA) via CRD	27/07/2021	n/a	22
International Network of Agencies for Health Technology Assessment (INAHTA)	27/07/2021	n/a	57
MEDLINE (Ovid)	27/07/2021	1946 to July Week 3 2021	5416
MEDLINE In-Process (Ovid)	27/07/2021	1946 to July 26, 2021	268
MEDLINE Epub Ahead of Print (Ovid)	27/07/2021	July 26, 2021	210
Total after deduplication			**15013**

Database: Ovid MEDLINE(R) <1946 to July 26, 2021>

1 exp Prostatic Neoplasms/ 134914
2 Prostatic Intraepithelial Neoplasia/ 1378
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 138662
4 (PCa or PrCa).tw. 38718
5 or/1-4 187916
6 *Risk Assessment/ 31987
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 319552
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 306
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 1971
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 93
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 1476
12 or/6-11 341357
13 5 and 12 6928
14 ((D'Amico or DAmico) adj6 prostat*).tw. 112
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 29
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 89
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 0
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 242
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 155
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 62
21 or/14-20 683
22 5 and 21 554
23 incidence.sh. 278986
24 exp mortality/403182
25 follow-up studies.sh. 667387
26 prognos:.tw. 560035
27 predict:.tw. 1418300
28 course:.tw. 570777
29 scor:.tw. 905529
30 observ:.mp. 3240349
31 or/23-30 6456805
32 22 and 31 349
33 13 or 32 7145
34 Animals/ not Humans/ 4831675
35 33 not 34 7086
36 limit 35 to english language 6747
37 limit 36 to ed=20070101-20210727 5571
38 Comment/ or Letter/ or Editorial/ or Historical article/ or (conference abstract or conference paper or "conference review" or letter or editorial or case report).pt. 2158312
39 37 not 38 5416

Database: Ovid MEDLINE(R) In-Process & In-Data-Review Citations <1946 to July 26, 2021>

Ovid MEDLINE(R) In-Process & In-Data-Review Citations <1946 to July 26, 2021>

1 exp Prostatic Neoplasms/ 0
2 Prostatic Intraepithelial Neoplasia/ 0
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 3346
4 (PCa or PrCa).tw. 1215
5 or/1-4 3940
6 *Risk Assessment/ 0
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 11390
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 12
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 51
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 4
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 49
12 or/6-11 11489
13 5 and 12 267
14 ((D'Amico or DAmico) adj6 prostat*).tw. 2
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 1
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 2
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 0
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 6
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 6
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 1
21 or/14-20 17
22 5 and 21 15
23 incidence.sh. 0
24 exp mortality/0
25 follow-up studies.sh. 0
26 prognos:.tw. 17864
27 predict:.tw. 43722
28 course:.tw. 8663
29 scor:.tw. 32076
30 observ:.mp. 61350
31 or/23-30 131235
32 22 and 31 6
33 13 or 32 270
34 Animals/ not Humans/ 0
35 33 not 34 270

36	limit 35 to english language	268
37	limit 36 to dt=20070101-20210727	268

Database: Ovid MEDLINE(R) Epub Ahead of Print <July 26, 2021>

1	exp Prostatic Neoplasms/	0
2	Prostatic Intraepithelial Neoplasia/	0
3	(prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw.	2610
4	(PCa or PrCa).tw.	985
5	or/1-4	3203
6	*Risk Assessment/	0
7	(risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw.	9879
8	(5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw.	10
9	(3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw.	57
10	(("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw.	1
11	("Cancer of the Prostate Risk Assessment" or CAPRA).tw.	22
12	or/6-11	9958
13	5 and 12	205
14	((D'Amico or DAmico) adj6 prostat*).tw.	1
15	(((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw.	1
16	(("European Association of Urology" or EAU) adj6 prostat*).tw.	2
17	(("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw.	0
18	(("American Urological Association" or AUA) adj6 prostat*).tw.	2
19	(("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw.	4
20	(("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw.	3
21	or/14-20	13
22	5 and 21	12
23	incidence.sh.	0
24	exp mortality/	0
25	follow-up studies.sh.	0
26	prognos:.tw.	11796
27	predict:.tw.	36308
28	course:.tw.	8668
29	scor:.tw.	28963
30	observ:.mp.	51723
31	or/23-30	111731
32	22 and 31	6
33	13 or 32	210
34	Animals/ not Humans/	0

35	33 not 34	210
36	limit 35 to english language	210

Database: Embase <1974 to 2021 July 26>

1	exp prostate tumor/	258155
2	prostatic intraepithelial neoplasia/	2932
3	(prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw.	234991
4	(PCa or PrCa).tw.	71165
5	or/1-4	335371
6	*risk assessment/	62073
7	(risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw.	557037
8	(5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw.	560
9	(3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw.	3272
10	(("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw.	150
11	("Cancer of the Prostate Risk Assessment" or CAPRA).tw.	2091
12	or/6-11	590549
13	5 and 12	15173
14	((D'Amico or DAmico) adj6 prostat*).tw.	379
15	(((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw.	79
16	(("European Association of Urology" or EAU) adj6 prostat*).tw.	296
17	(("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw.	3
18	(("American Urological Association" or AUA) adj6 prostat*).tw.	786
19	(("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw.	473
20	(("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw.	117
21	or/14-20	2111
22	5 and 21	1778
23	incidence.sh.	460937
24	exp mortality/	1169932
25	follow-up.sh.	1712512
26	prognos:.tw.	1000308
27	predict:.tw.	2328169
28	course:.tw.	879614
29	scor:.tw.	1689446
30	observ:.mp.	4836589
31	or/23-30	10611820
32	22 and 31	1323
33	13 or 32	16028
34	Nonhuman/ not Human/	4827852

35 33 not 34 15891
36 limit 35 to english language 15481
37 limit 36 to dc=20070101-20210727 14149
38 Comment/ or Letter/ or Editorial/ or Historical article/ or (conference abstract or conference paper or "conference review" or letter or editorial or case report).pt. 6826849
39 37 not 38 7604

Database: Cochrane Database of Systematic Reviews (CDSR) and Cochrane Central Register of Controlled Trials (CENTRAL)

#1 MeSH descriptor: [Prostatic Neoplasms] explode all trees 5746
#2 MeSH descriptor: [Prostatic Intraepithelial Neoplasia] this term only 47
#3 (prostat* near/4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)):ti,ab,kw 15280
#4 (PCa or PrCa):ti,ab,kw 5247
#5 {or #1-#4} 19446
#6 MeSH descriptor: [Risk Assessment] this term only 9027
#7 (risk* near/2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)):ti,ab,kw 58202
#8 (5 NEXT tier* or "5tier*" or five NEXT tier* or "5-strata*" or "5strata*" or "five-strata*"):ti,ab,kw 27
#9 (3 NEXT tier* or "3tier*" or three NEXT tier* or "3-strata*" or "3strata*" or "three-strata*"):ti,ab,kw 230
#10 (("Cambridge Prognostic Group*" or CPG*) near/6 prostat*):ti,ab,kw 1
#11 ("Cancer of the Prostate Risk Assessment" or CAPRA):ti,ab,kw 52
#12 {or #6-#11} 58463
#13 #5 and #12 1394
#14 ((D'Amico or DAmico) near/6 prostat*):ti,ab,kw 12
#15 (((National Institute near/4 Excellence) or NICE) near/6 prostat*):ti,ab,kw 319
#16 (("European Association of Urology" or EAU) near/6 prostat*):ti,ab,kw 19
#17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) near/6 prostat*):ti,ab,kw 0
#18 (("American Urological Association" or AUA) near/6 prostat*):ti,ab,kw 67
#19 (("National Comprehensive Cancer Network" or NCCN) near/6 prostat*):ti,ab,kw 36
#20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) near/6 prostat*):ti,ab,kw 13
#21 {or #14-#20} 456
#22 #5 and #21 250
#23 #13 or #22 with Publication Year from 2007 to 2021, with Cochrane Library publication date Between Jan 2007 and Jul 2021, in Trials 1346

Database: Database of Abstracts of Reviews of Effect (DARE) and Health Technology Assessment (HTA)

1	MeSH DESCRIPTOR Prostatic Neoplasms EXPLODE ALL TREES	709
2	MeSH DESCRIPTOR Prostatic Intraepithelial Neoplasia EXPLODE ALL TREES	2
3	(prostat* near (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*))	912
4	(PCa or PrCa)	44
5	#1 OR #2 OR #3 OR #4	956
6	MeSH DESCRIPTOR Risk Assessment EXPLODE ALL TREES	2129
7	(risk* near (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*))	7398
8	(("5-tier*" or "5tier*" or "five-tier*" or "5-strata*" or "5strata*" or "five-strata*"))	9
9	(("3-tier*" or "3tier*" or "three-tier*" or "3-strata*" or "3strata*" or "three-strata*"))	10
10	((("Cambridge Prognostic Group*" or CPG*) near prostat*))	0
11	(("Cancer of the Prostate Risk Assessment" or CAPRA))	17
12	#6 OR #7 OR #8 OR #9 OR #10 OR #11	7426
13	#5 AND #12	158
14	(((D'Amico or DAmico) near prostat*))	0
15	((((National Institute near/4 Excellence) or NICE) near prostat*))	1
16	((("European Association of Urology" or EAU) near prostat*))	0
17	((("Genito-Urinary Radiation Oncologists of Canada" or GUROC) near prostat*))	0
18	((("American Urological Association" or AUA) near prostat*))	7
19	((("National Comprehensive Cancer Network" or NCCN) near prostat*))	0
20	((("Memorial Sloan Kettering Cancer Center" or MSKCC) near prostat*))	0
21	#14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20	8
22	#5 AND #21	3
23	#13 OR #22	161
24	* FROM 2007 TO 2021	56435
25	#23 AND #24	120
26	(#23 and #24) IN DARE FROM 2007 TO 2021	90
27	(#23 and #24) IN HTA FROM 2007 TO 2021	22

Database: International Network of Agencies for Health Technology Assessment

25 #24 AND #23 57
24 * FROM 2007 TO 2021 11822
23 #22 OR #13 79
22 #21 AND #5 11
21 #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 20
20 ("Memorial Sloan Kettering Cancer Center" or MSKCC) AND (prostat*) 0

19 ("National Comprehensive Cancer Network" or NCCN) AND (prostat*) 1

18 ("American Urological Association" or AUA) AND (prostat*) 3
17 ("Genito-Urinary Radiation Oncologists of Canada" or GUROC) AND (prostat*) 0
16 (("European Association of Urology" or EAU)) AND (prostat*) 0
15 ((National Institute near Excellence) or NICE) AND (prostat*) 16
14 (D'Amico or DAmico) AND (prostat*) 0
13 #12 AND #5 70
12 #11 OR #10 OR #9 OR #8 OR #7 OR #6 2627
11 ("Cancer of the Prostate Risk Assessment" or CAPRA) 1
10 ("Cambridge Prognostic Group*" or CPG*) AND (prostat*) 0
9 ("3-tier*" or "3tier*" or "three-tier*" or "3-strata*" or "3strata*" or "three-strata*") 14
8 ("5-tier*" or "5tier*" or "five-tier*" or "5-strata*" or "5strata*" or "five-strata*") 14
7 (risk*) AND (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*) 2584
6 "Risk Assessment"[mh] 120
5 #4 OR #3 OR #2 OR #1 337
4 PCa or PrCa 4
3 (prostat*) AND (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*) 330
2 "Prostatic Intraepithelial Neoplasia"[mh] 0
1 "Prostatic Diseases"[mh] 3

Cost effectiveness searches

Databases	Date searched	Version/files	No. of results downloaded
EconLit (Ovid)	28/07/2021	1886 to July 22, 2021	18
Embase (Ovid) (apply economics filter)	28/07/2021	1974 to 2021 July 27	53

NHS Economic Evaluation Database (NHS EED) via CRD	28/07/2021	n/a	8
International Network of Agencies for Health Technology Assessment (INAHTA)	28/07/2021	n/a	57
MEDLINE (Ovid) (apply economics filter)	28/07/2021	1946 to July Week 3 2021	60
MEDLINE In-Process (Ovid) (apply economics filter)	28/07/2021	1946 to July 27, 2021	1
MEDLINE Epub Ahead of Print (apply economics filter)	28/07/2021	July 27, 2021	3
Total			**200**

Database: Ovid MEDLINE(R) <1946 to July Week 3 2021>

1 exp Prostatic Neoplasms/ 134708
2 Prostatic Intraepithelial Neoplasia/ 1377
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 138421
4 (PCa or PrCa).tw. 38619
5 or/1-4 187595
6 *Risk Assessment/ 31943
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 318877
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 305
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 1967
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 93
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 1471
12 or/6-11 340660
13 5 and 12 6904
14 ((D'Amico or DAmico) adj6 prostat*).tw. 112
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 29
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 89
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 0
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 242
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 155
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 62
21 or/14-20 683
22 5 and 21 554

23 13 or 22 7294
24 Cost-Benefit Analysis/ 85425
25 (cost* and ((qualit* adj2 adjust* adj2 life*) or qaly*)).tw. 12147
26 ((incremental* adj2 cost*) or ICER).tw. 12527
27 (cost adj2 utilit*).tw. 4809
28 (cost* and ((net adj benefit*) or (net adj monetary adj benefit*) or (net adj health adj benefit*))).tw. 1552
29 ((cost adj2 (effect* or utilit*)) and (quality adj of adj life)).tw. 16701
30 (cost and (effect* or utilit*)).ti. 28683
31 or/24-30 96490
32 23 and 31 79
33 limit 32 to english language 77
34 Comment/ or Letter/ or Editorial/ or Historical article/ or (conference abstract or conference paper or "conference review" or letter or editorial or case report).pt. 2156129
35 33 not 34 75
36 limit 35 to ed=20070101-20210728 60

Database: Ovid MEDLINE(R) In-Process & In-Data-Review Citations <1946 to July 27, 2021>

1 exp Prostatic Neoplasms/ 0
2 Prostatic Intraepithelial Neoplasia/ 0
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 3319
4 (PCa or PrCa).tw. 1204
5 or/1-4 3907
6 *Risk Assessment/ 0
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 11341
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 11
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 52
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 4
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 49
12 or/6-11 11439
13 5 and 12 269
14 ((D'Amico or DAmico) adj6 prostat*).tw. 2
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 1
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 2
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 0
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 6
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 6
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 1

21 or/14-20 17
22 5 and 21 15
23 13 or 22 279
24 Cost-Benefit Analysis/ 0
25 (cost* and ((qualit* adj2 adjust* adj2 life*) or qaly*)).tw. 549
26 ((incremental* adj2 cost*) or ICER).tw. 554
27 (cost adj2 utilit*).tw. 181
28 (cost* and ((net adj benefit*) or (net adj monetary adj benefit*) or (net adj health adj benefit*))).tw. 75
29 ((cost adj2 (effect* or utilit*)) and (quality adj of adj life)).tw. 654
30 (cost and (effect* or utilit*)).ti. 729
31 or/24-30 1199
32 23 and 31 1
33 limit 32 to english language 1

Database: Ovid MEDLINE(R) Epub Ahead of Print <July 27, 2021>

1 exp Prostatic Neoplasms/ 0
2 Prostatic Intraepithelial Neoplasia/ 0
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 2606
4 (PCa or PrCa).tw. 986
5 or/1-4 3201
6 *Risk Assessment/ 0
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 9854
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 10
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 55
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 1
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 22
12 or/6-11 9932
13 5 and 12 204
14 ((D'Amico or DAmico) adj6 prostat*).tw. 1
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 1
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 2
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 0
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 2
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 4
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 3
21 or/14-20 13
22 5 and 21 12
23 13 or 22 213

24 Cost-Benefit Analysis/ 0
25 (cost* and ((qualit* adj2 adjust* adj2 life*) or qaly*)).tw. 451
26 ((incremental* adj2 cost*) or ICER).tw. 393
27 (cost adj2 utilit*).tw. 211
28 (cost* and ((net adj benefit*) or (net adj monetary adj benefit*) or (net adj health adj benefit*))).tw. 58
29 ((cost adj2 (effect* or utilit*)) and (quality adj of adj life)).tw. 614
30 (cost and (effect* or utilit*)).ti. 625
31 or/24-30 1206
32 23 and 31 3
33 limit 32 to english language 3

Database: Embase <1974 to 2021 July 27>

1 exp prostate tumor/ 258208
2 prostatic intraepithelial neoplasia/ 2932
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 235059
4 (PCa or PrCa).tw. 71187
5 or/1-4 335452
6 *risk assessment/ 62103
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 557218
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 560
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 3273
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 150
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 2091
12 or/6-11 590746
13 5 and 12 15179
14 ((D'Amico or DAmico) adj6 prostat*).tw. 379
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 79
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 296
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 3
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 786
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 473
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 117
21 or/14-20 2111
22 5 and 21 1778
23 13 or 22 16437
24 cost utility analysis/ 10510
25 (cost* and ((qualit* adj2 adjust* adj2 life*) or qaly*)).tw. 24946
26 ((incremental* adj2 cost*) or ICER).tw. 25554

27 (cost adj2 utilit*).tw. 9233
28 (cost* and ((net adj benefit*) or (net adj monetary adj benefit*) or (net adj health adj benefit*))).tw. 2581
29 ((cost adj2 (effect* or utilit*)) and (quality adj of adj life)).tw. 30465
30 (cost and (effect* or utilit*)).ti. 49557
31 or/24-30 78192
32 23 and 31 115
33 limit 32 to english language 112
34 Comment/ or Letter/ or Editorial/ or Historical article/ or (conference abstract or conference paper or "conference review" or letter or editorial or case report).pt. 6828217
35 33 not 34 61
36 limit 35 to dc=20070101-20210728 53

Database: Econlit <1886 to July 22, 2021>

1 exp Prostatic Neoplasms/ 0
2 Prostatic Intraepithelial Neoplasia/ 0
3 (prostat* adj4 (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)).tw. 109
4 (PCa or PrCa).tw. 488
5 or/1-4 593
6 *Risk Assessment/ 0
7 (risk* adj2 (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)).tw. 20186
8 (5-tier* or 5tier* or five-tier* or 5-strata* or 5strata* or five-strata*).tw. 8
9 (3-tier* or 3tier* or three-tier* or 3-strata* or 3strata* or three-strata*).tw. 196
10 (("Cambridge Prognostic Group*" or CPG*) adj6 prostat*).tw. 0
11 ("Cancer of the Prostate Risk Assessment" or CAPRA).tw. 9
12 or/6-11 20394
13 5 and 12 18
14 ((D'Amico or DAmico) adj6 prostat*).tw. 0
15 (((National Institute adj4 Excellence) or NICE) adj6 prostat*).tw. 2
16 (("European Association of Urology" or EAU) adj6 prostat*).tw. 0
17 (("Genito-Urinary Radiation Oncologists of Canada" or GUROC) adj6 prostat*).tw. 0
18 (("American Urological Association" or AUA) adj6 prostat*).tw. 0
19 (("National Comprehensive Cancer Network" or NCCN) adj6 prostat*).tw. 0
20 (("Memorial Sloan Kettering Cancer Center" or MSKCC) adj6 prostat*).tw. 0
21 or/14-20 2
22 5 and 21 0
23 13 or 22 18

Database: NHS Economic Evaluation Database (NHS EED)

1 MeSH DESCRIPTOR Prostatic Neoplasms EXPLODE ALL TREES 709
2 MeSH DESCRIPTOR Prostatic Intraepithelial Neoplasia EXPLODE ALL TREES 2
3 (prostat* near (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*)) 912
4 (PCa or PrCa) 44
5 #1 OR #2 OR #3 OR #4 956
6 MeSH DESCRIPTOR Risk Assessment EXPLODE ALL TREES 2129
7 (risk* near (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*)) 7398
8 (("5-tier*" or "5tier*" or "five-tier*" or "5-strata*" or "5strata*" or "five-strata*"))
9 (("3-tier*" or "3tier*" or "three-tier*" or "3-strata*" or "3strata*" or "three-strata*"))
10 ((("Cambridge Prognostic Group*" or CPG*) near prostat*)) 0
11 (("Cancer of the Prostate Risk Assessment" or CAPRA)) 17
12 #6 OR #7 OR #8 OR #9 OR #10 OR #11 7426
13 #5 AND #12 158
14 (((D'Amico or DAmico) near prostat*)) 0
15 ((((National Institute near/4 Excellence) or NICE) near prostat*)) 1
16 ((("European Association of Urology" or EAU) near prostat*)) 0
17 ((("Genito-Urinary Radiation Oncologists of Canada" or GUROC) near prostat*)) 0
18 ((("American Urological Association" or AUA) near prostat*)) 7
19 ((("National Comprehensive Cancer Network" or NCCN) near prostat*)) 0
20 ((("Memorial Sloan Kettering Cancer Center" or MSKCC) near prostat*)) 0
21 #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 8
22 #5 AND #21 3
23 #13 OR #22 161
24 * FROM 2007 TO 2021 56435
25 #23 AND #24 120
26 (#23 and #24) IN DARE FROM 2007 TO 2021 90
27 (#23 and #24) IN HTA FROM 2007 TO 2021 22
28 (#23 and #24) IN NHSEED FROM 2007 TO 2021 8

Database: International Network of Agencies for Health Technology Assessment

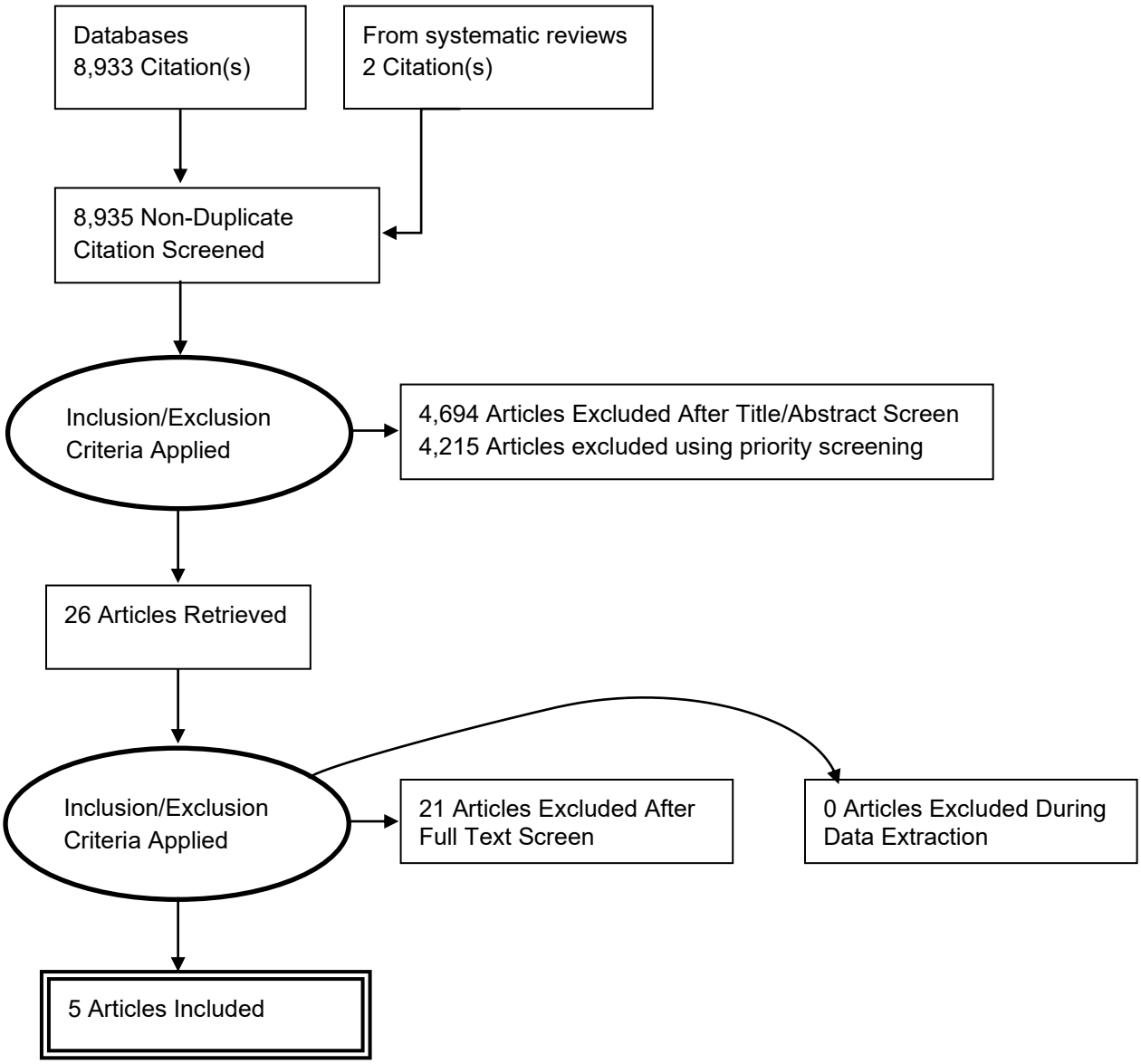
25 #24 AND #23 57
24 * FROM 2007 TO 2021 11822
23 #22 OR #13 79

22 #21 AND #5 11
21 #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 20
20 ("Memorial Sloan Kettering Cancer Center" or MSKCC) AND (prostat*) 0

19 ("National Comprehensive Cancer Network" or NCCN) AND (prostat*) 1

18 ("American Urological Association" or AUA) AND (prostat*) 3
17 ("Genito-Urinary Radiation Oncologists of Canada" or GUROC) AND (prostat*) 0
16 (("European Association of Urology" or EAU)) AND (prostat*) 0
15 ((National Institute near Excellence) or NICE) AND (prostat*) 16
14 (D'Amico or DAmico) AND (prostat*) 0
13 #12 AND #5 70
12 #11 OR #10 OR #9 OR #8 OR #7 OR #6 2627
11 ("Cancer of the Prostate Risk Assessment" or CAPRA) 1
10 ("Cambridge Prognostic Group*" or CPG*) AND (prostat*) 0
9 ("3-tier*" or "3tier*" or "three-tier*" or "3-strata*" or "3strata*" or "three-strata*") 14
8 ("5-tier*" or "5tier*" or "five-tier*" or "5-strata*" or "5strata*" or "five-strata*") 14
7 (risk*) AND (stratif* or assess* or analy* or benefit* or classifi* or model* or tool* or adjust* or evaluat* or categor* or system* or score* or level* or check* or group* or grade*) 2584
6 "Risk Assessment"[mh] 120
5 #4 OR #3 OR #2 OR #1 337
4 PCa or PrCa 4
3 (prostat*) AND (neoplas* or cancer* or carcinoma* or adenocarcinom* or tumour* or tumor* or malignan* or metasta* or angiosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or carcino* or leiomyosarcoma* or lump* or disease*) 330
2 "Prostatic Intraepithelial Neoplasia"[mh] 0
1 "Prostatic Diseases"[mh] 3

Appendix C – Prognostic evidence study selection



Appendix D – Prognostic evidence

Abdel-Rahman, 2018

Bibliographic Reference	Abdel-Rahman, Omar; Dissecting the heterogeneity of localized prostate cancer risk groups through integration of percent of positive cores.; Future oncology (London, England); 2018; vol. 14 (no. 15); 1469-1476

Study Characteristics

Study design	Retrospective cohort study
Study details	Study location US Study setting The Surveillance, Epidemiology, and End Results (SEER) database was used to identify eligible clinically localised prostate adenocarcinoma patients Study dates 2010 to 2014
Inclusion criteria	Criteria 1 Patients with N0/M0 disease according to the TNM sixth system Criteria 2 Those who were not treated with radical surgery
Exclusion criteria	Criteria 1 Cases with no information about T stage, number of examined cores, and number of positive cores, PSA level or Gleason

	score Criteria 2 Cases with less than six scores examined
Number of participants and recruitment methods	Validation cohort (n=30,445) from the SEER database; eligible participants were identified using the ICD-O-3/WHO 2008 category of 'prostate'
Length of follow-up	Median follow-up for all participants was 27 months (range: 1 to 59 months)
Loss to follow up	
Outcome(s) of interest	C-statistic using prostate cancer specific mortality as the dependent variable among the validation cohort
Prognostic factors or risk factor(s) or sign(s)/symptom(s)	D'Amico risk stratification model: • low risk • intermediate risk • high risk Modified risk stratification model (incorporation of percent of positive cores into the D'Amico risk stratification model); • low risk • intermediate risk 1: core positive ratio ≤50% • intermediate risk 2: core positive ratio >50% • high risk 1: core positive ratio ≤50% • high risk 2: core positive ratio >50%
Additional comments	

Population characteristics
Study-level characteristics

Characteristic	Study (N = 30445)
Age groups	
Less than 70	n = 18882 ; % = 62
Sample size	
70 or more	n = 11563 ; % = 38
Sample size	
Ethnicity	
White	n = 22455 ; % = 73.8
Sample size	
Black	n = 5615 ; % = 18.4
Sample size	
Others	n = 1602 ; % = 5.3
Sample size	
Unknown	n = 773 ; % = 2.5
Sample size	
Histology	
Adenocarcinoma, not otherwise specified	n = 30341 ; % = 99.7

Characteristic	Study (N = 30445)
Sample size	
Other variants Sample size	n = 104 ; % = 0.3
Grade group **1** Sample size	n = 14188 ; % = 46.6
2 Sample size	n = 7650 ; % = 25.1
3 Sample size	n = 3706 ; % = 12.2
4 Sample size	n = 2840 ; % = 9.3
5 Sample size	n = 2061 ; % = 6.8
Gleason score **3** Sample size	n = 5 ; % = 0.01
4	n = 16 ; % = 0.1

Characteristic	Study (N = 30445)
Sample size	
5 Sample size	n = 56 ; % = 0.2
6 Sample size	n = 14111 ; % = 46.3
7 Sample size	n = 11356 ; % = 37.3
8 Sample size	n = 2840 ; % = 9.3
9 Sample size	n = 1853 ; % = 6.1
10 Sample size	n = 208 ; % = 0.7
PSA level **>10** Sample size	n = 22570 ; % = 74.1
10 to 20 Sample size	n = 5062 ; % = 16.6

Characteristic	Study (N = 30445)
≤20 Sample size	n = 2813 ; % = 9.3
T stage **T1 to T2a** Sample size	n = 27409 ; % = 90
T2b Sample size	n = 1033 ; % = 3.4
T2c to T3 Sample size	n = 1902 ; % = 6.3
t4 Sample size	n = 101 ; % = 0.3
Examined cores Custom value	Mean 12.48 (range 6 to 100)
Positive cores Custom value	Mean 4.21 (range 0 to 91)
Core positive ratio **50% or less** Sample size	n = 23708 ; % = 77.9

Characteristic	Study (N = 30445)
More than 50% Sample size	n = 6737 ; % = 22.1
Risk groups **Low** Sample size	n = 11809 ; % = 38.8
Intermediate Sample size	n = 11481 ; % = 37.7
High Sample size	n = 7155 ; % = 23.5
AJCC sixth stages **I** Sample size	n = 7 ; % = 0.01
II Sample size	n = 29568 ; % = 97.1
III Sample size	n = 769 ; % = 2.5
IV Sample size	n = 101 ; % = 0.3

Characteristic	Study (N = 30445)
Radiotherapy **Yes** Sample size	n = 18243 ; % = 59.9
No/unknown Sample size	n = 12202 ; % = 40.1
Chemotherapy **Yes** Sample size	n = 64 ; % = 0.2
No/unknown Sample size	n = 30381 ; % = 99.8

Validation cohort

Critical appraisal - GUT PROBAST tool

Section	Question	Answer
Selection of participants	Overall risk of bias for selection of participants domain	Low
Selection of participants	Concerns for applicability for selection of participants domain	Low

Section	Question	Answer
Predictors or their assessment	Overall risk of bias for predictors or their assessment domain	Unclear *(No information on whether predictors were assessed without knowledge of outcome data.)*
Predictors or their assessment	Concerns for applicability for predictors or their assessment domain	Low
Outcome or its determination	Overall risk of bias for outcome or its determination domain	Unclear *(No information on whether outcome was determined without knowledge of predictor.)*
Outcome or its determination	Concerns for applicability for outcome or its determination domain	Low
Analysis	Overall risk of bias for analysis domain	High *(No information about the use of competing risks analysis in predicting prostate cancer-specific mortality.)*
Overall Risk of bias and Applicability	Risk of bias	High
Overall Risk of bias and Applicability	Concerns for applicability	Low

Gnanapragasam, 2018

Bibliographic Reference Gnanapragasam, V J; Bratt, O; Muir, K; Lee, L S; Huang, H H; Stattin, P; Lophatananon, A; The Cambridge Prognostic Groups for improved prediction of disease mortality at diagnosis in primary non-metastatic prostate cancer: a validation study.; BMC medicine; 2018; vol. 16 (no. 1); 31

Study Characteristics

Study design	Retrospective cohort study
Study details	Study location Sweden Singapore Study setting The Prostate Cancer data Base Sweden. Cohort database from the Singapore Health Study. Study dates Swedish cohort was followed until 31 December 2015. No information on dates for the cohort from Singapore. Sources of funding There was no specific funding for the study.

Inclusion criteria	Criteria 1 No evidence of metastatic disease (Mx or M0) Criteria 2 PSA <100 ng/ml
Exclusion criteria	Criteria 1 Lack of data in PSA, clinical T stage or Gleason Grade Group
Number of participants and recruitment methods	Sweden cohort (n=72,337) Singapore cohort (n=2,550)
Length of follow-up	Sweden cohort (median 7 years) Singapore cohort (median 4.1 years)
Loss to follow up	
Outcome(s) of interest	Prostate cancer specific mortality (Cox proportional hazards regression model and the log rank test with pair-wise comparisons were used; “Low risk” was the reference group in the NICE model and “CPG1” in the CPGroup model). Concordance index (c-statistic) was used for model discrimination (sub-hazard ratios were used in computation instead of hazard ratio to account for competing risks from other causes of death).
Prognostic factors or risk factor(s) or sign(s)/symptom(s)	NICE risk stratification model: • low risk • intermediate risk • high risk Cambridge Prognostic Group criteria:

	1. Gleason score 6 (Grade Group 1) **AND** PSA <10 ng/ml **AND** Stages T1–T2 2. Gleason score 3 + 4 = 7 (Grade Group 2) **OR** PSA 10–20 ng/ml **AND** Stages T1–T2 3. Gleason score 3 + 4 = 7 (Grade Group 2) **AND** PSA 10–20 ng/ml **AND** Stages T1–T2 **OR** Gleason 4 + 3 = 7 (Grade Group 3) **AND** Stages T1–T2 4. One of Gleason score 8 (Grade Group 4) **OR** PSA > 20 ng/ml **OR** Stage T3 5. Any combination of Gleason score 8 (Grade Group 4), PSA > 20 ng/ml or Stage T3 **OR** Gleason score 9–10 (Grade Group 5) **OR** Stage T4
Covariates adjusted for in the multivariable regression modelling	

Study arms

Sweden cohort (N = 72337)

Singapore cohort (N = 2550)

Population characteristics

Arm-level characteristics

Characteristic	Sweden cohort (N = 72337)	Singapore cohort (N = 2550)
Age groups **Less than 60** Sample size	n = 10309	n = 501

Characteristic	Sweden cohort (N = 72337)	Singapore cohort (N = 2550)
60 to 69 Sample size	n = 28903	n = 1198
70 to 79 Sample size	n = 23483	n = 739
80 or more Sample size	n = 9642	n = 112
Gleason score **Grade group 1 <6** Sample size	n = 39572	n = 1127
Grade group 2: 3+4 Sample size	n = 14112	n = 723
Grade group 3: 4+3 Sample size	n = 7892	n = 327
Grade group 4: 8 Sample size	n = 6527	n = 170
Grade group 5: 9 to 10 Sample size	n = 4234	n = 203
PSA level		

Characteristic	Sweden cohort (N = 72337)	Singapore cohort (N = 2550)
Less than 10 Sample size	n = 38690	n = 1344
10 to 20 Sample size	n = 18357	n = 682
More than 20 Sample size	n = 15290	n = 524
T stage **T1** Sample size	n = 37270	n = 1626
T2 Sample size	n = 23473	n = 661
T3 Sample size	n = 10825	n = 246
T4 Sample size	n = 769	n = 17

Critical appraisal - GUT PROBAST tool

Section	Question	Answer
Selection of participants	Overall risk of bias for selection of participants domain	Low
Selection of participants	Concerns for applicability for selection of participants domain	Low
Predictors or their assessment	Overall risk of bias for predictors or their assessment domain	Unclear *(No information on whether predictors were assessed without knowledge of outcome data.)*
Predictors or their assessment	Concerns for applicability for predictors or their assessment domain	Low
Outcome or its determination	Overall risk of bias for outcome or its determination domain	Unclear *(No information on whether outcome was determined without knowledge of predictor.)*
Outcome or its determination	Concerns for applicability for outcome or its determination domain	Low
Analysis	Overall risk of bias for analysis domain	Low
Overall Risk of bias and Applicability	Risk of bias	Moderate
Overall Risk of bias and Applicability	Concerns for applicability	Low

Gnanapragasam, 2016

Bibliographic Reference Gnanapragasam, Vincent J; Lophatananon, Artitaya; Wright, Karen A; Muir, Kenneth R; Gavin, Anna; Greenberg, David C; Improving Clinical Risk Stratification at Diagnosis in Primary Prostate Cancer: A Prognostic Modelling Study.; PLoS medicine; 2016; vol. 13 (no. 8); e1002063

Study Characteristics

Study design	Retrospective cohort study
Study details	Study location UK Study setting Northern Ireland Cancer registry Study dates 1 January 2000 to 30 September 2013 Sources of funding There was no specific funding for the project.
Inclusion criteria	Criteria 1 Only cases with all components of diagnostic stage, primary and secondary grade, and presenting PSA (ng/ml) as well as data on follow-up and survival were included as these variables were essential to build the risk model
Exclusion criteria	Criteria 1 Cases with any metastatic involvement (as documented by M stage disease and/or positive bone or CT scan)

Number of participants and recruitment methods	Validation cohort (n=1,706) from an independent dataset from the Northern Ireland Cancer Registry, which has information on all population PSA tests linked to prostate cancer diagnosis and death
Length of follow-up	Median 4.8 years
Outcome(s) of interest	For model discrimination, concordance index (c-statistic) was used with inclusion of competing risks for prostate-cancer-specific mortality
Prognostic factors or risk factor(s) or sign(s)/symptom(s)	NICE risk stratification model: • low risk • intermediate risk • high risk Cambridge Prognostic Group criteria: 1. Gleason 6 (prognostic score 1) **AND** PSA <10 ng/ml **AND** Stage T1–T2 2. Gleason 3 + 4 = 7 (prognostic score 2) **OR** PSA 10–20 ng/ml **AND** Stage T1–T2 3. Gleason 3 + 4 = 7 (prognostic score 2) **AND** PSA 10–20 ng/ml **AND** Stage T1–T2 **OR** Gleason 4 + 3 = 7 (prognostic score 3) **AND** Stage T1–T2 4. Any one of Gleason 8 (prognostic score 4) **OR** PSA > 20 ng/ml **OR** Stage T3 5. More than one of Gleason 8 (prognostic score 4), PSA > 20 ng/ml, Stage T3 **OR** Any Gleason 9 to 10 (prognostic score 5) **OR** Any Stage T4
Covariates adjusted for in the multivariable regression modelling	

Population characteristics

Study-level characteristics

Characteristic	Study (N = 1706)
Age groups **Less than 60** Sample size	n = 321
60 to 69 Sample size	n = 723
70 to 79 Sample size	n = 559
80 or more Sample size	n = 103
Gleason score **Prognostic score 1: <6** Sample size	n = 587
Prognostic score 2: 3+4 Sample size	n = 487
Prognostic score 3: 4+3 Sample size	n = 210

Characteristic	Study (N = 1706)
Prognostic score 4: 8 Sample size	n = 192
Prognostic score 5: 9 to 10 Sample size	n = 230
PSA level **Less than 10** Sample size	n = 711
10 to 20 Sample size	n = 589
More than 20 Sample size	n = 406
T stage **T1** Sample size	n = 585
T2 Sample size	n = 578
T3 Sample size	n = 537

Characteristic	Study (N = 1706)
T4 Sample size	n = 6

Critical appraisal - GUT PROBAST tool

Section	Question	Answer
Selection of participants	Overall risk of bias for selection of participants domain	Low
Selection of participants	Concerns for applicability for selection of participants domain	Low
Predictors or their assessment	Overall risk of bias for predictors or their assessment domain	Unclear *(No information on whether predictors were assessed without knowledge of outcome data.)*
Predictors or their assessment	Concerns for applicability for predictors or their assessment domain	Low
Outcome or its determination	Overall risk of bias for outcome or its determination domain	Unclear *(No information on whether outcome was determined without knowledge of predictor.)*
Outcome or its determination	Concerns for applicability for outcome or its determination domain	Low
Analysis	Overall risk of bias for analysis domain	Low

Section	Question	Answer
Overall Risk of bias and Applicability	Risk of bias	Moderate
Overall Risk of bias and Applicability	Concerns for applicability	Low

Lee, 2021

Bibliographic Reference Lee, Changhee; Light, Alexander; Alaa, Ahmed; Thurtle, David; van der Schaar, Mihaela; Gnanapragasam, Vincent J; Application of a novel machine learning framework for predicting non-metastatic prostate cancer-specific mortality in men using the Surveillance, Epidemiology, and End Results (SEER) database.; The Lancet. Digital health; 2021; vol. 3 (no. 3); e158-e165

Study Characteristics

Study design	Retrospective cohort study
Study details	Study location US Study setting Data was collected from the Surveillance, Epidemiology, and End Results (SEER) Program. Study dates January 1st 2000 to December 31th 2016 Sources of funding There was no funding source for the study.
Inclusion criteria	Criteria 1 Men aged 35 to 95 years diagnosed with histologically confirmed non-metastatic prostate cancer (site code C61.9)
Exclusion criteria	Criteria 1 Evidence of metastatic disease (including lymph node metastasis)

	Criteria 2 Those with missing survival data or data on PSA, Gleason grade, or stage Criteria 3 Men younger than 35 years or older than 95 years
Number of participants and recruitment methods	Participants (n=171,942) were identified from the SEER database using the site code C61.9. The SEER cohort was randomly split (64:16:20) into the training, validation, or testing sets.
Length of follow-up	Median 6.1 years
Outcome(s) of interest	Model discrimination was assessed using the concordance index (c-index or c-statistic) for predicting 10-year prostate cancer specific mortality; calibration was assessed using Brier scores.
Prognostic factors or risk factor(s) or sign(s)/symptom(s)	**Tier based risk stratification models with the variables incorporated to each model:** • Cancer of the Prostate Risk Assessment score (age, PSA, biopsy core involvement, T stage, Gleason grade or Grade Group) • Cambridge Prognostic Groups (PSA, T stage, Gleason grade or Grade Group) • National Comprehensive Cancer Care Network (PSA, biopsy core involvement, T stage, Gleason grade or Grade Group) • Genitourinary Radiation Oncologists of Canada (PSA, T stage, Gleason grade or Grade Group) • American Urological Association (PSA, T stage, Gleason grade or Grade Group) • European Association of Urology (PSA, T stage, Gleason grade or Grade Group) • National Institute for Health and Care Excellence (PSA, T stage, Gleason grade or Grade Group)
Additional comments	Data was only shown for the testing set (c-index and Brier scores) but it was reported in the article that the c-index for predicting prostate cancer-specific mortality was consistently high in training, validation, and testing sets with good calibration.

Population characteristics
Study-level characteristics

Characteristic	Study (N = 171942)
Age (years) Mean (SD)	65.6 (8.9)
Ethnicity **White** Sample size	n = 134139 ; % = 78
Black Sample size	n = 24488 ; % = 14.2
Asian Sample size	n = 8962 ; % = 5.2
PSA level (ng/mL) Mean (SD)	10.1 (13.3)
T stage **T1a** Sample size	n = 1811 ; % = 1.1
T1b Sample size	n = 1026 ; % = 0.6

Characteristic	Study (N = 171942)
T1c Sample size	n = 101036 ; % = 58.8
T2a Sample size	n = 48690 ; % = 28.3
T2b Sample size	n = 11282 ; % = 6.6
T2c Sample size	n = 4728 ; % = 2.8
T3a Sample size	n = 1699 ; % = 1
T3b Sample size	n = 1195 ; % = 0.7
T4 Sample size	n = 475 ; % = 0.3
Core involvement Data available for 66885 (38.9%) of 171942 men **Cores taken** Mean (SD)	12.4 (2.5)

Characteristic	Study (N = 171942)
Cores positive Mean (SD)	4.2 (2.1)
Cores negative Mean (SD)	8.1 (2.7)
Primary Gleason score **2** Sample size	n = 9 ; % = 0
3 Sample size	n = 126083 ; % = 73.3
4 Sample size	n = 42588 ; % = 24.8
5 Sample size	n = 3262 ; % = 1.9
Secondary Gleason score **2** Sample size	n = 6 ; % = 0
3 Sample size	n = 94715 ; % = 55.1

Characteristic	Study (N = 171942)
4 Sample size	n = 67284 ; % = 39.1
5 Sample size	n = 9937 ; % = 5.7
Grade group **1** Sample size	n = 72548 ; % = 42.2
2 Sample size	n = 52245 ; % = 30.4
3 Sample size	n = 21086 ; % = 12.7
4 Sample size	n = 14675 ; % = 8.5
5 Sample size	n = 10668 ; % = 6.2

Critical appraisal - GUT PROBAST tool

Section	Question	Answer
Selection of participants	Overall risk of bias for selection of participants domain	Low
Selection of participants	Concerns for applicability for selection of participants domain	Low
Predictors or their assessment	Overall risk of bias for predictors or their assessment domain	Unclear *(No information on whether predictors were assessed without knowledge of outcome data.)*
Predictors or their assessment	Concerns for applicability for predictors or their assessment domain	Low
Outcome or its determination	Overall risk of bias for outcome or its determination domain	Unclear *(No information on whether outcome was determined without knowledge of predictor.)*
Outcome or its determination	Concerns for applicability for outcome or its determination domain	Low
Analysis	Overall risk of bias for analysis domain	High *(Mean imputation was used to handle missing data; no information about the use of competing risks analysis in predicting prostate cancer-specific mortality.)*
Overall Risk of bias and Applicability	Risk of bias	High
Overall Risk of bias and Applicability	Concerns for applicability	Low

Zelic, 2020

Bibliographic Reference Zelic, Renata; Garmo, Hans; Zugna, Daniela; Stattin, Par; Richiardi, Lorenzo; Akre, Olof; Pettersson, Andreas; Predicting Prostate Cancer Death with Different Pretreatment Risk Stratification Tools: A Head-to-head Comparison in a Nationwide Cohort Study.; European urology; 2020; vol. 77 (no. 2); 180-188

Study Characteristics

Study design	Retrospective cohort study
Study details	Study location Sweden Study setting The Prostate Cancer data Base Sweden. Study dates January 1, 1998 to December 31, 2016 Sources of funding Work was supported by the Strategic Research Programme in Cancer (StratCan) and the Strategic Research Program in Epidemiology Young Scholar Award (AP) at Karolinska Institute, the Swedish Cancer Society (2011/825), and the Stockholm County Council.
Inclusion criteria	Criteria 1 Men diagnosed with non-metastatic (not M1 or N1) prostate cancer
Number of participants and	n=139,515

recruitment methods	
Length of follow-up	Median 5.83 years
Loss to follow up	
Outcome(s) of interest	Prostate cancer specific mortality, defined as prostate cancer listed as the underlying cause of death (ICD-10 code: C61). Cause-specific hazards for prostate cancer death and death from other causes were combined to obtain cumulative incidence functions (CIFs) for prostate cancer death. Discrimination was evaluated by concordance index (C statistic) adapted for competing risks.
Prognostic factors or risk factor(s) or sign(s)/symptom(s)	Tier-based risk stratification tools: • AUA = American Urological Association • CAPRA = Cancer of the Prostate Risk Assessment • CPG = Cambridge Prognostic Groups • EAU= European Association of Urology • GUROC = Genito-Urinary Radiation Oncologists of Canada • NCCN = National Comprehensive Cancer Network • NICE = The National Institute for Health and Care Excellence
Covariates adjusted for in the multivariable regression modelling	
Additional comments	Information on the individual biopsy cores was not available in Prostate Cancer data Base Sweden version 4. Therefore, core level information could not be used in the construction of the risk groups for AUA, AUA-i, and NCCN risk stratification models.

Population characteristics
Study-level characteristics

Characteristic	Study (N = 139515)
Age at diagnosis (years) Median (IQR)	69 (63 to 76)
PSA (ng/mL) Median (IQR)	9.4 (5.8 to 20)
Prostate volume (ml) Median (IQR)	38 (29 to 52)
Clinical tumour stage **T1** Sample size	n = 65804 ; % = 49.37
T1a Sample size	n = 5426 ; % = 7.27
T1b Sample size	n = 3493 ; % = 4.68
T1c Sample size	n = 65682 ; % = 88.04
T2 Sample size	n = 48444 ; % = 35.61

Characteristic	Study (N = 139515)
T3a Sample size	n = 21726 ; % = 16.02
Biopsy Gleason score Missing data (n=14,639) **6 or less** Sample size	n = 60546 ; % = 47.08
7 Sample size	n = 47215 ; % = 36.71
3 + 4 Sample size	n = 28680 ; % = 65.95
4 + 3 Sample size	n = 14810 ; % = 34.05
8 Sample size	n = 11559 ; % = 8.99
9 Sample size	n = 8552 ; % = 6.65
>10 Sample size	n = 729 ; % = 0.57

Characteristic	Study (N = 139515)
Primary Gleason grade Missing data (n=22,602) **1** Sample size	n = 112 ; % = 0.1
2 Sample size	n = 3706 ; % = 3.17
3 Sample size	n = 80229 ; % = 68.62
4 Sample size	n = 30237 ; % = 25.86
5 Sample size	n = 2629 ; % = 2.25
Secondary Gleason grade Missing data (n=22,776) **1** Sample size	n = 31 ; % = 0.03
2 Sample size	n = 3517 ; % = 3.01

Characteristic	Study (N = 139515)
3 Sample size	n = 65608 ; % = 56.2
4 Sample size	n = 39704 ; % = 34.01
5 Sample size	n = 7879 ; % = 6.75
Number of cores sampled at biopsy Missing data (n=44,118) Median (IQR)	10 (8 to 12)
Total length of biopsy cores (mm) Missing data (n=83,258) Median (IQR)	146 (119 to 172)
Number of cores with cancer Missing data (n=44,826) Median (IQR)	3 (2 to 5)
Total length of cancer (mm) Missing data (n=77,667) Median (IQR)	9.4 (3 to 26)

Critical appraisal - GUT PROBAST tool

Section	Question	Answer
Selection of participants	Overall risk of bias for selection of participants domain	Low
Selection of participants	Concerns for applicability for selection of participants domain	Low
Predictors or their assessment	Overall risk of bias for predictors or their assessment domain	Unclear *(No information on whether predictors were assessed without knowledge of outcome data.)*
Predictors or their assessment	Concerns for applicability for predictors or their assessment domain	Low
Outcome or its determination	Overall risk of bias for outcome or its determination domain	Unclear *(No information on whether outcome was determined without knowledge of predictors.)*
Outcome or its determination	Concerns for applicability for outcome or its determination domain	Low
Analysis	Overall risk of bias for analysis domain	Low
Overall Risk of bias and Applicability	Risk of bias	Moderate
Overall Risk of bias and Applicability	Concerns for applicability	Low

Appendix E – Forest plots

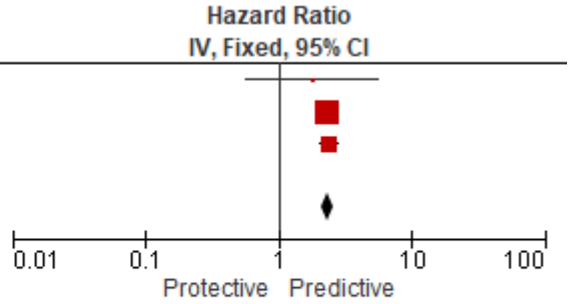
Hazard ratios

5 tier prostate cancer risk stratification models

Figure 1: Cambridge Prognostic Group risk stratification model for prediction of prostate cancer specific mortality (higher HR means CPG2 is predictive of prostate cancer specific mortality)

Study or Subgroup	log[Hazard Ratio]	SE	CPG1 Total	CPG2 Total	Weight	Hazard Ratio IV, Fixed, 95% CI
Gnanapragasam 2018 - Singapore cohort	0.5822	0.5839	734	621	0.7%	1.79 [0.57, 5.62]
Gnanapragasam 2018 - Sweden cohort	0.8329	0.0612	25303	14796	62.1%	2.30 [2.04, 2.59]
Zelic 2020 (1)	0.8629	0.079	0	0	37.2%	2.37 [2.03, 2.77]
Total (95% CI)			**26037**	**15417**	**100.0%**	**2.32 [2.11, 2.55]**

Heterogeneity: Chi² = 0.29, df = 2 (P = 0.87); I² = 0%
Test for overall effect: Z = 17.47 (P < 0.00001)

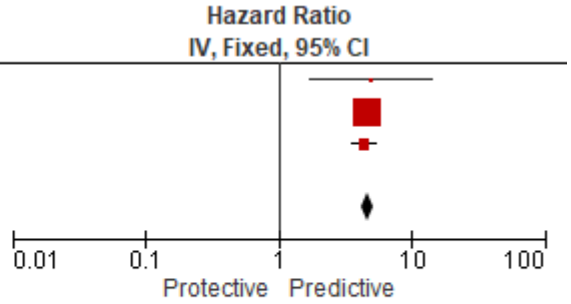


Footnotes
(1) Zelic 2020 did not report number of participants per group (total number of participants n=139,515)

Figure 2: Cambridge Prognostic Group risk stratification model for prediction of prostate cancer specific mortality (higher HR means CPG3 is predictive of prostate cancer specific mortality)

Study or Subgroup	log[Hazard Ratio]	SE	CPG1 Total	CPG3 Total	Weight	Hazard Ratio IV, Fixed, 95% CI
Gnanapragasam 2018 - Singapore cohort	1.5872	0.5391	734	386	1.0%	4.89 [1.70, 14.07]
Gnanapragasam 2018 - Sweden cohort	1.5476	0.061	25303	7354	74.6%	4.70 [4.17, 5.30]
Zelic 2020 (1)	1.4816	0.1067	0	0	24.4%	4.40 [3.57, 5.42]
Total (95% CI)			**26037**	**7740**	**100.0%**	**4.63 [4.17, 5.13]**

Heterogeneity: Chi² = 0.30, df = 2 (P = 0.86); I² = 0%
Test for overall effect: Z = 29.07 (P < 0.00001)

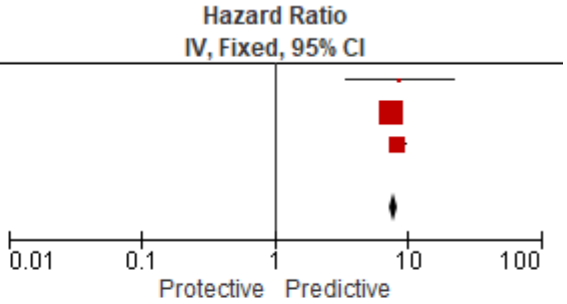


Footnotes
(1) Zelic 2020 did not report number of participants per group (total number of participants n=139,515)

Figure 3: Cambridge Prognostic Group risk stratification model for prediction of prostate cancer specific mortality (higher HR means CPG4 is predictive of prostate cancer specific mortality)

Study or Subgroup	log[Hazard Ratio]	SE	CPG1 Total	CPG4 Total	Weight	Hazard Ratio IV, Fixed, 95% CI
Gnanapragasam 2018 - Singapore cohort	2.1691	0.4793	734	430	0.7%	8.75 [3.42, 22.39]
Gnanapragasam 2018 - Sweden cohort	2.0042	0.0513	25303	13506	61.3%	7.42 [6.71, 8.21]
Zelic 2020 (1)	2.1304	0.0652	0	0	38.0%	8.42 [7.41, 9.57]
Total (95% CI)			**26037**	**13936**	**100.0%**	**7.79 [7.20, 8.43]**

Heterogeneity: Chi² = 2.37, df = 2 (P = 0.31); I² = 16%
Test for overall effect: Z = 51.11 (P < 0.00001)

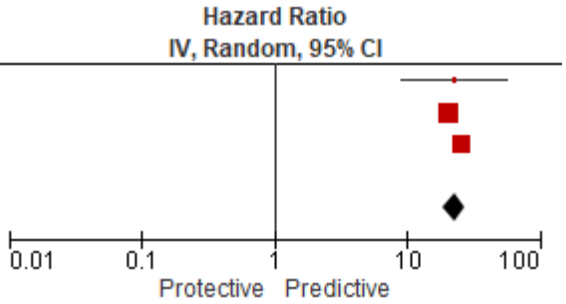


Footnotes
(1) Zelic 2020 did not report number of participants per group (total number of participants n=139,515)

Figure 4: Cambridge Prognostic Group risk stratification model for prediction of prostate cancer specific mortality (higher HR means CPG5 is predictive of prostate cancer specific mortality)

Study or Subgroup	log[Hazard Ratio]	SE	CPG1 Total	CPG5 Total	Weight	Hazard Ratio IV, Random, 95% CI
Gnanapragasam 2018 - Singapore cohort	3.1259	0.4671	734	379	3.9%	22.78 [9.12, 56.91]
Gnanapragasam 2018 - Sweden cohort	3.0214	0.0485	25303	11378	50.3%	20.52 [18.66, 22.57]
Zelic 2020 (1)	3.2349	0.0643	0	0	45.8%	25.40 [22.40, 28.82]
Total (95% CI)			**26037**	**11757**	**100.0%**	**22.72 [18.83, 27.42]**

Heterogeneity: Tau² = 0.02; Chi² = 7.03, df = 2 (P = 0.03); I² = 72%
Test for overall effect: Z = 32.56 (P < 0.00001)

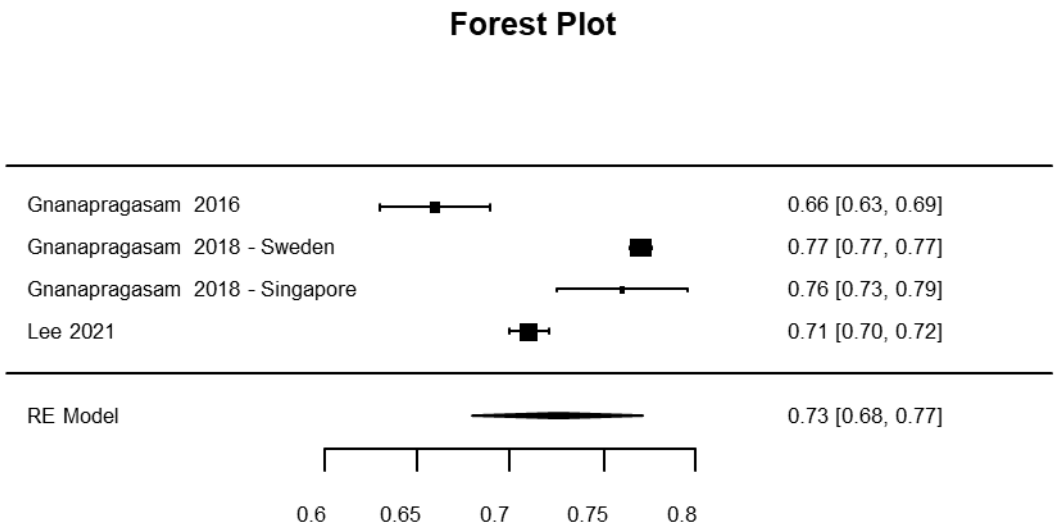


Footnotes
(1) Zelic 2020 did not report number of participants per group (total number of participants n=139,515)

C-statistics

3 tier prostate cancer risk stratification models

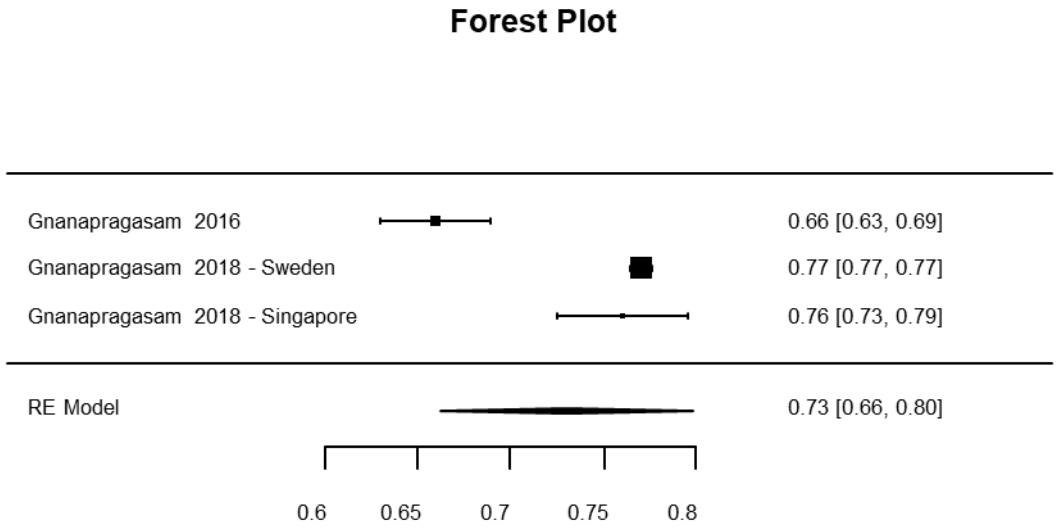
Figure 5: NICE risk stratification model for prediction of prostate cancer specific mortality



Study	Weight
Gnanapragasam 2016	23.93%
Gnanapragasam 2018 Sweden cohort	26.63%
Gnanapragasam 2018 Singapore cohort	23.06%
Lee 2021	26.38%

RE model, I^2 = 98.01%

Figure 6: NICE risk stratification model for prediction of prostate cancer specific mortality – sensitivity analysis without studies at high risk of bias

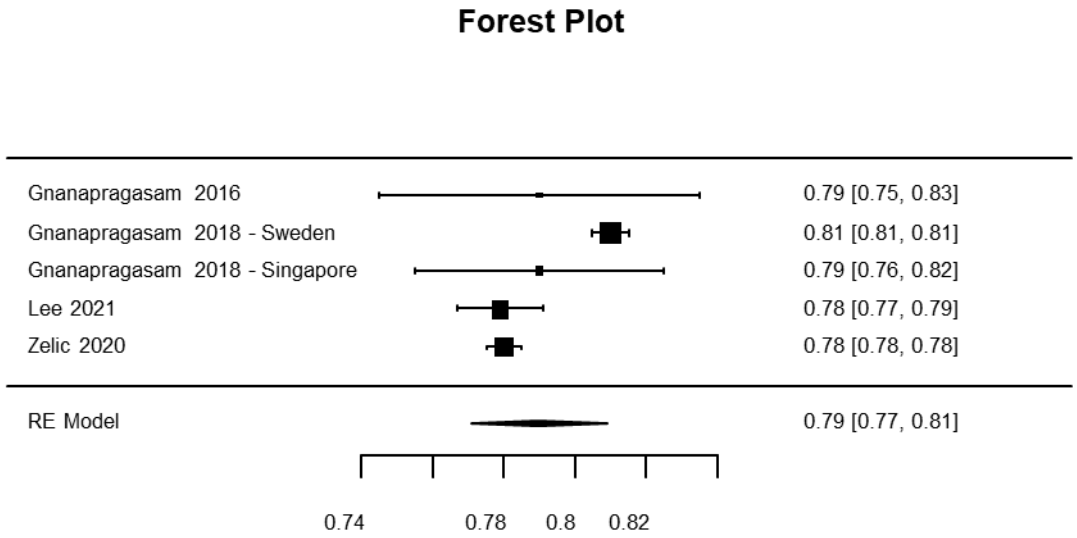


Study	Weight
Gnanapragasam 2016	32.85%
Gnanapragasam 2018 Sweden cohort	35.05%
Gnanapragasam 2018 Singapore cohort	32.10%

RE model, I^2 = 96.03%

5 tier prostate cancer risk stratification models

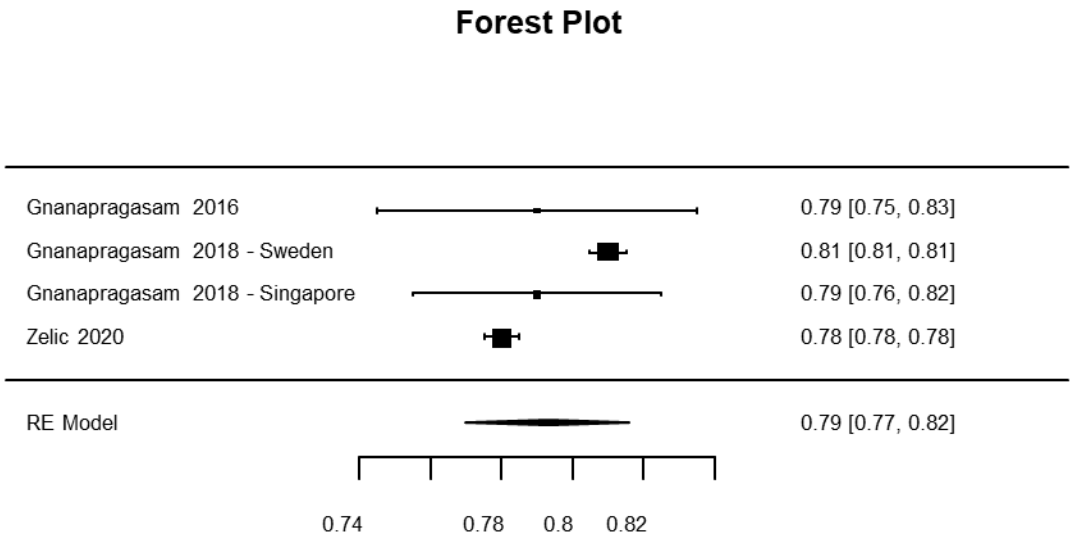
Figure 7: Cambridge Prognostic Group risk stratification model for prediction of prostate cancer specific mortality



Study	Weight
Gnanapragasam 2016	10.68%
Gnanapragasam 2018 Sweden cohort	25.79%
Gnanapragasam 2018 Singapore cohort	13.95%
Lee 2021	23.79%
Zelic 2020	25.79%

RE model, I^2 = 94.70%

Figure 8: Cambridge Prognostic Group risk stratification model for prediction of prostate cancer specific mortality – sensitivity analysis without studies at high risk of bias



Study	Weight
Gnanapragasam 2016	14.76%
Gnanapragasam 2018 Sweden cohort	33.14%
Gnanapragasam 2018 Singapore cohort	18.96%
Zelic 2020	33.14%

RE model, I^2 = 95.67%

Appendix F – GRADE tables

Prostate cancer specific mortality

Hazard ratios

3 tier prostate cancer risk stratification models

No. of studies	Study design	No. of participants		Hazard ratio (95% CI)	Absolute effect (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
		Comparator	Reference							
NICE risk stratification model for prediction of prostate cancer specific mortality Higher HR means intermediate risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		2.94 (2.51, 3.44)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
NICE risk stratification model for prediction of prostate cancer specific mortality Higher HR means high risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		14.16 (12.42, 16.14)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
D'Amico risk stratification model for prediction of prostate cancer specific mortality Higher HR means intermediate risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		2.88 (2.45, 3.38)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
D'Amico risk stratification model for prediction of prostate cancer specific mortality Higher HR means high risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		13.69 (12.00, 15.62)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
EAU risk stratification model for prediction of prostate cancer specific mortality Higher HR means intermediate risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		2.94 (2.51, 3.44)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate

No. of studies	Study design	No. of participants		Hazard ratio (95% CI)	Absolute effect (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
		Comparator	Reference							
EAU risk stratification model for prediction of prostate cancer specific mortality Higher HR means high risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		14.16 (12.42, 16.14)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
GUROC risk stratification model for prediction of prostate cancer specific mortality Higher HR means intermediate risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		3.22 (2.77, 3.76)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
GUROC risk stratification model for prediction of prostate cancer specific mortality Higher HR means high risk is predictive of prostate cancer specific mortality (reference: low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		16.08 (14.10, 18.35)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate

a. Study did not report number of participants for comparator and reference groups
b. >33.3% of weighted data from studies at moderate or high risk of bias

European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence), N/A (not applicable, single study), N/C (not calculable)

5 tier prostate cancer risk stratification models

No. of studies	Study design	No. of participants		Hazard ratio (95% CI)	Absolute effect (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
		Comparator	Reference							
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG2 is predictive of prostate cancer specific mortality (reference: CPG1)										
Gnanapragasam 2018 Singapore cohort Sweden cohort	Retrospective cohort	621	734	2.32 (2.11, 2.55)	25 more per 1000 (20 more to 30 more)	Serious[b]	Not serious	Not serious	Not serious	Moderate
		14,796	25,303							
		Total sample 139,515 [a]								

No. of studies	Study design	No. of participants		Hazard ratio (95% CI)	Absolute effect (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
		Comparator	Reference							
Zelic 2020										
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG3 is predictive of prostate cancer specific mortality (reference: CPG1)										
Gnanapragasam 2018 Singapore cohort Sweden cohort Zelic 2020	Retrospective cohort	386	734	4.63 (4.17, 5.13)	69 more per 1000 (60 more to 78 more)	Serious[b]	Not serious	Not serious	Not serious	Moderate
		7,354	25,303							
		Total sample 139,515 [a]								
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG4 is predictive of prostate cancer specific mortality (reference: CPG1)										
Gnanapragasam 2018 Singapore cohort Sweden cohort Zelic 2020	Retrospective cohort	430	734	7.79 (7.20, 8.43)	129 more per 1000 (118 more to 141 more)	Serious[b]	Not serious	Not serious	Not serious	Moderate
		13,506	25,303							
		Total sample 139,515 [a]								
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG5 is predictive of prostate cancer specific mortality (reference: CPG1)										
Gnanapragasam 2018 Singapore cohort Sweden cohort	Retrospective cohort	379	734	22.72 (18.83, 27.42)	413 more per 1000 (339 more to 502 more)	Serious[b]	Not serious	Very serious[c]	Not serious	Very low
		11,378	25,303							
		Total sample 139,515 [a]								

No. of studies	Study design	No. of participants		Hazard ratio (95% CI)	Absolute effect (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
		Comparator	Reference							
Zelic 2020										
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG2 is predictive of prostate cancer specific mortality (reference: CPG1)										
Gnanapragasam 2018 Sweden cohort	Retrospective cohort	14,796	25,303	2.30 (2.04, 2.59)	251 more per 1000 (201 more to 307 more)	Serious[b]	Not serious	N/A	Not serious	Moderate
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG3 is predictive of prostate cancer specific mortality (reference: CPG2)										
Gnanapragasam 2018 Sweden cohort	Retrospective cohort	7,354	14,796	2.11 (1.89, 2.36)	47 more per 1000 (37 more to 57 more)	Serious[b]	Not serious	N/A	Not serious	Moderate
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG4 is predictive of prostate cancer specific mortality (reference: CPG3)										
Gnanapragasam 2018 Sweden cohort	Retrospective cohort	13,506	7,354	1.56 (1.42, 1.72)	45 more per 1000 (34 more to 58 more)	Serious[b]	Not serious	N/A	Not serious	Moderate
CPG risk stratification model for prediction of prostate cancer specific mortality Higher HR means CPG5 is predictive of prostate cancer specific mortality (reference: CPG4)										
Gnanapragasam 2018 Sweden cohort	Retrospective cohort	11,378	13,506	2.72 (2.58, 2.88)	234 more per 1000 (215 more to 256 more)	Serious[b]	Not serious	N/A	Not serious	Moderate

No. of studies	Study design	No. of participants		Hazard ratio (95% CI)	Absolute effect (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
		Comparator	Reference							
AUA-i risk stratification model for prediction of prostate cancer specific mortality Higher HR means low risk is predictive of prostate cancer specific mortality (reference: very low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		1.11 (0.83, 1.49)	N/C	Serious[b]	Not serious	N/A	Serious[d]	Low
AUA-i risk stratification model for prediction of prostate cancer specific mortality Higher HR means favourable intermediate risk is predictive of prostate cancer specific mortality (reference: very low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		2.54 (2.00, 3.23)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
AUA-i risk stratification model for prediction of prostate cancer specific mortality Higher HR means unfavourable intermediate risk is predictive of prostate cancer specific mortality (reference: very low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		5.15 (4.05, 6.55)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate
AUA-i risk stratification model for prediction of prostate cancer specific mortality Higher HR means high risk is predictive of prostate cancer specific mortality (reference: very low risk)										
Zelic 2020	Retrospective cohort	Total sample 139,515 [a]		17.64 (14.12, 22.05)	N/C	Serious[b]	Not serious	N/A	Not serious	Moderate

a. Study did not report number of participants for comparator and reference groups
b. >33.3% of weighted data from studies at moderate or high risk of bias
c. i-squared >66.7%
d. 95% confidence interval crosses the line of no effect

American Urological Association with favourable and non-favourable intermediate groups (AUA-i), Cambridge Prognostic Groups (CPG), N/A (not applicable, single study), N/C (not calculable)

C-statistic

3 tier prostate cancer risk stratification models

No. of studies	Study design	Sample size	C-statistic (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
NICE risk stratification model for prediction of prostate cancer specific mortality, median 5.9 years follow-up								

No. of studies	Study design	Sample size	C-statistic (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort Lee 2021	Retrospective cohort	248,535	0.73 (0.68, 0.77)	Serious[a]	Not serious	Very serious[b]	Serious[c]	Very low
NICE risk stratification model for prediction of prostate cancer specific mortality, median 4.8 years follow-up – sensitivity analysis without studies at high risk of bias								
Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort	Retrospective cohort	76,593	0.73 (0.66, 0.80)	Serious[a]	Not serious	Very serious[b]	Serious[c]	Very low
NICE risk stratification model for prediction of prostate cancer specific mortality, 10 years follow-up								
Zelic 2020	Retrospective cohort	139,515	0.73*	Serious[a]	Not serious	N/A	Not serious	Moderate
D'Amico risk stratification model for prediction of prostate cancer specific mortality, 10 years follow-up								
Zelic 2020	Retrospective cohort	139,515	0.73 (0.72, 0.73)	Serious[a]	Not serious	N/A	Not serious	Moderate
D'Amico risk stratification modified model for prediction of prostate cancer specific mortality, median 2.25 years follow-up								
Abdel-Rahman 2018	Retrospective cohort	30,445	0.78 (0.75, 0.81)	Very serious[d]	Not serious	N/A	Serious[c]	Very low
EAU risk stratification model for prediction of prostate cancer specific mortality, 10 years follow-up								
Lee 2021	Retrospective cohort	171,942	0.71 (070, 0.72)	Very serious[d]	Not serious	N/A	Not serious	Low
GUROC risk stratification model for prediction of prostate cancer specific mortality, 10 years follow-up								

No. of studies	Study design	Sample size	C-statistic (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
Lee 2021	Retrospective cohort	171,942	0.75 (0.73, 0.76)	Very serious[d]	Not serious	N/A	Not serious	Low

a. >33.3% of weighted data from studies at moderate or high risk of bias
b. i-squared >66.7%
c. 95% confidence interval crosses 2 categories of test classification accuracy
d. >33.3% of weighted data from studies at high risk of bias
* 95% confidence interval not provided or calculable
European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence), N/A (not applicable, single study)

5 tier prostate cancer risk stratification models

No. of studies	Study design	Sample size	C-statistic (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
CPG risk stratification model for prediction of prostate cancer specific mortality, median 7 years follow-up								
Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort Sweden cohort Lee 2021 Zelic 2020	Retrospective cohort	388,050	0.79 (0.77, 0.81)	Serious[a]	Not serious	Very serious[b]	Serious[c]	Very low
CPG risk stratification model for prediction of prostate cancer specific mortality, median 5.9 years follow-up – sensitivity analysis without studies at high risk of bias								
Gnanapragasam 2016 Gnanapragasam 2018 Singapore cohort	Retrospective cohort	216,108	0.79 (0.77, 0.82)	Serious[a]	Not serious	Very serious[b]	Serious[c]	Very low

No. of studies	Study design	Sample size	C-statistic (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
Sweden cohort Zelic 2020								
D'Amico risk stratification modified model for prediction of prostate cancer specific mortality, median 2.25 years follow-up								
Abdel-Rahman 2018	Retrospective cohort	30,445	0.81 (0.78, 0.84)	Very serious[d]	Not serious	N/A	Serious[c]	Very low

a. >33.3% of weighted data from studies at moderate or high risk of bias
b. i-squared >66.7%
c. 95% confidence interval crosses 2 categories of test classification accuracy
d. >33.3% of weighted data from studies at high risk of bias
* 95% confidence interval not provided or calculable
Cambridge Prognostic Groups (CPG), N/A (not applicable, single study)

Brier score

3 tier prostate cancer risk stratification models

No. of studies	Study design	Sample size	Brier score (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
NICE risk stratification model for prediction of prostate cancer specific mortality								
Lee 2021	Retrospective cohort	171,942	0.039 (0.037, 0.041)	Very serious[a]	Not serious	N/A	No serious	Low
EAU risk stratification model for prediction of prostate cancer specific mortality								
Lee 2021	Retrospective cohort	171,942	0.039 (0.037, 0.041)	Very serious[a]	Not serious	N/A	No serious	Low
GUROC risk stratification model for prediction of prostate cancer specific mortality								
Lee 2021	Retrospective cohort	171,942	0.039 (0.037, 0.041)	Very serious[a]	Not serious	N/A	No serious	Low

a. Study at high risk of bias
European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), NICE (National Institute for Health and Care Excellence), N/A (not applicable, single study), N/C (not calculable)

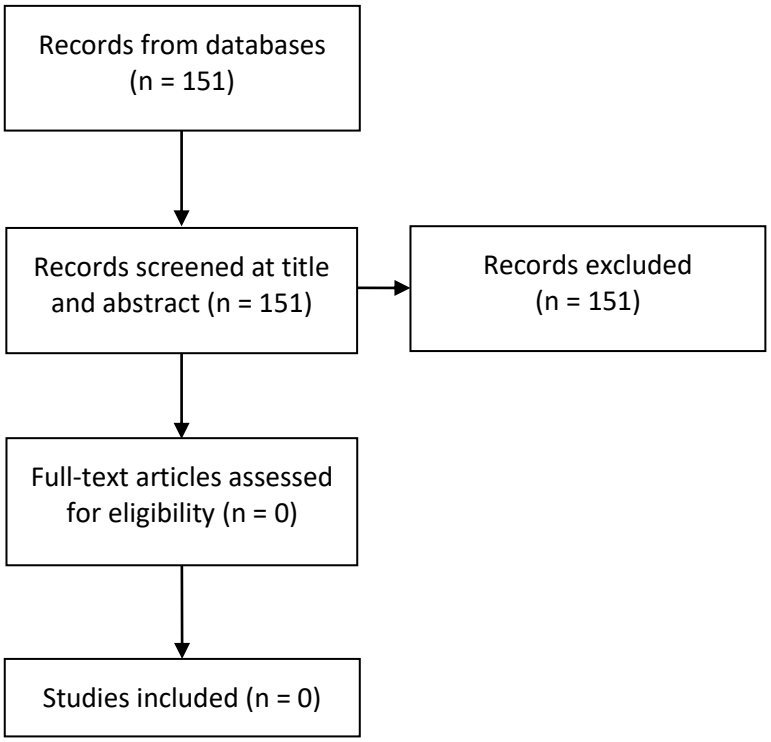
5 tier prostate cancer risk stratification models

No. of studies	Study design	Sample size	Brier score (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
CPG risk stratification model for prediction of prostate cancer specific mortality								
Lee 2021	Retrospective cohort	171,942	0.037 (0.035, 0.039)	Very serious[a]	Not serious	N/A	No serious	Low

a. Study at high risk of bias

Cambridge Prognostic Groups (CPG), N/A (not applicable, single study), N/C (not calculable)

Appendix G – Economic evidence study selection



Appendix H – Economic evidence tables

No economic evidence was identified for this review question.

Appendix I – Health economic model

No health economic modelling was conducted for this review question.

Appendix J – Excluded studies

Prognostic evidence

Study	Reason for exclusion
Algohary, Ahmad, Shiradkar, Rakesh, Pahwa, Shivani et al. (2020) Combination of peri-tumoral and intra-tumoral radiomic features on bi-parametric mri accurately stratifies prostate cancer risk: A multi-site study. Cancers 12(8): 1-14	- Outcome to be predicted do not match that specified in the protocol
Briganti, Alberto, Passoni, Niccolo, Ferrari, Matteo et al. (2010) When to perform bone scan in patients with newly diagnosed prostate cancer: external validation of the currently available guidelines and proposal of a novel risk stratification tool. European urology 57(4): 551-8	- End point do not match that specified in the protocol Bone scan was done at diagnosis
Chun, Felix K-H, Karakiewicz, Pierre I, Briganti, Alberto et al. (2007) A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. BJU international 99(4): 794-800	- Review article but not a systematic review
Feuer EJ, Lee M, Mariotto AB et al. (2012) The Cancer Survival Query System: making survival estimates from the Surveillance, Epidemiology, and End Results program more timely and relevant for recently diagnosed patients. Cancer 118(22): 5652-5662	- Outcome to be predicted do not match that specified in the protocol Life expectancy
Feuer, Eric J, Rabin, Borsika A, Zou, Zhaohui et al. (2014) The Surveillance, Epidemiology, and End Results Cancer Survival Calculator SEER*CSC: validation in a managed care setting. Journal of the National Cancer Institute. Monographs 2014(49): 265-74	- Population does not match that specified in the protocol Participants were not newly diagnosed
Gnanapragasam, Vincent J, Barrett, Tristan, Thankapannair, Vineetha et al. (2019) Using prognosis to guide inclusion criteria, define standardised endpoints and stratify follow-up in active surveillance for prostate cancer. BJU international 124(5): 758-767	- Outcome to be predicted do not match that specified in the protocol Effects of treatment in the different CPG groups
Hiremath, Amogh, Shiradkar, Rakesh, Fu, Pingfu et al. (2021) An integrated nomogram combining deep learning, Prostate Imaging-Reporting and Data System (PI-RADS) scoring, and clinical variables for identification of clinically significant prostate cancer on biparametric MRI: a retrospective multicentre study. The Lancet Digital Health 3(7): e445-e454	- Assessment tool do not match that specified in the protocol
Howlader N, Mariotto AB, Woloshin S et al. (2014) Providing clinicians and patients with actual prognosis: cancer in the context of competing causes of death. Journal of the National Cancer Institute. Monographs 2014(49): 255-264	- Assessment tool do not match that specified in the protocol
Izumi, Kouji, Ikeda, Hiroko, Maolake, Aerken et al. (2015) The relationship between prostate-specific antigen and TNM classification or Gleason score in prostate cancer patients with low prostate-specific antigen levels. The Prostate 75(10): 1034-42	- Assessment tool do not match that specified in the protocol
Lorent, Marine, Maalmi, Haifa, Tessier, Philippe et al. (2019) Meta-analysis of predictive models to assess the	- Population does not match that specified in the protocol

Study	Reason for exclusion
clinical validity and utility for patient-centered medical decision making: application to the CAncer of the Prostate Risk Assessment (CAPRA). BMC medical informatics and decision making 19(1): 2	Participants underwent radical prostatectomy
Parry, M G, Cowling, T E, Sujenthiran, A et al. (2020) Risk stratification for prostate cancer management: value of the Cambridge Prognostic Group classification for assessing treatment allocation. BMC medicine 18(1): 114	- Outcome to be predicted do not match that specified in the protocol Disease treatment
Rodrigues, George, Lukka, Himu, Warde, Padraig et al. (2013) The prostate cancer risk stratification (ProCaRS) project: recursive partitioning risk stratification analysis. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 109(2): 204-10	- Population does not match that specified in the protocol Radiotherapy patients
Rogasch, Julian M., Amthauer, Holger, Furth, Christian et al. (2018) Ga-68-PSMA PET/CT in treatment-naive patients with prostate cancer: Which clinical parameters and risk stratification systems best predict PSMA-positive metastases?. Prostate 78(14): 1103-1110	- End point do not match that specified in the protocol Metastases was identified at initial staging examination
Thurtle, David R, Greenberg, David C, Lee, Lui S et al. (2019) Individual prognosis at diagnosis in nonmetastatic prostate cancer: Development and external validation of the PREDICT Prostate multivariable model. PLoS medicine 16(3): e1002758	- Population does not match that specified in the protocol Participants were not newly diagnosed
Thurtle, David, Bratt, Ola, Stattin, Par et al. (2020) Comparative performance and external validation of the multivariable PREDICT Prostate tool for non-metastatic prostate cancer: a study in 69,206 men from Prostate Cancer data Base Sweden (PCBaSe). BMC medicine 18(1): 139	- Population does not match that specified in the protocol Participants were not newly diagnosed
Thurtle, David, Rossi, Sabrina H, Berry, Brendan et al. (2019) Models predicting survival to guide treatment decision-making in newly diagnosed primary non-metastatic prostate cancer: a systematic review. BMJ open 9(6): e029149	- Systematic review used as a reference for individual studies
Varghese, Bino, Chen, Frank, Hwang, Darryl et al. (2019) Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images. Scientific reports 9(1): 1570	- Assessment tool do not match that specified in the protocol
Xiao, Wen-Jun, Zhu, Yu, Zhu, Yao et al. (2018) Evaluation of clinical staging of the American Joint Committee on Cancer (eighth edition) for prostate cancer. World journal of urology 36(5): 769-774	- Assessment tool do not match that specified in the protocol
Xie, Mu, Gao, Xian-Shu, Ma, Ming-Wei et al. (2021) Population-Based Comparison of Different Risk Stratification Systems Among Prostate Cancer Patients. Frontiers in Oncology 11: 646073	- Population does not match that specified in the protocol Participants were not newly diagnosed
Yoshioka, Yasuo and Inoue, Takehiro (2007) Prostate Risk Index (PRIX) as a new method of risk classification for clinically localized prostate cancer. Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft ... [et al] 183(9): 490-6	- Model development study without validation data
Zelic, Renata, Pettersson, Andreas, Garmo, Hans et al. (2020) Corrigendum re "Predicting Prostate Cancer Death with Different Pretreatment Risk Stratification	- Erratum of Zelic 2020

Study	Reason for exclusion
Tools: A Head-to-head Comparison in a Nationwide Cohort Study" [Eur Urol 2020;77:180-8](S0302283819307559)(10.1016/j.eururo.2019.09.027). European Urology 78(1): e45-e47	

Appendix K – Methods

K.1 Selecting studies for inclusion

All references identified by the literature searches and from other sources (for example, from published systematic review) were uploaded into EPPI reviewer software (version 5) and de-duplicated. Titles and abstracts were assessed for possible inclusion using the criteria specified in the review protocol. 10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer.

This evidence review made use of the priority screening functionality within the EPPI-reviewer software. This functionality uses a machine learning algorithm (specifically, an SGD classifier) to take information on features (1, 2 and 3 word blocks) in the titles and abstract of papers marked as being 'includes' or 'excludes' during the title and abstract screening process, and re-orders the remaining records from most likely to least likely to be an include, based on that algorithm. This re-ordering of the remaining records occurs every time 25 additional records have been screened. Research is currently ongoing as to what are the appropriate thresholds where reviewing of abstracts can be stopped, assuming a defined threshold for the proportion of relevant papers it is acceptable to miss on primary screening. As a conservative approach until that research has been completed, the following rules were adopted during the production of this guideline:

- In every review, at least 50% of the identified abstracts (or 1,000 records, if that is a greater number) were always screened.
- After this point, screening was only terminated if a pre-specified threshold was met for a number of abstracts being screened without a single new include being identified. This threshold was set according to the expected proportion of includes in the review (with reviews with a lower proportion of includes needing a higher number of papers without an identified study to justify termination) and was always a minimum of 250.

As an additional check to ensure this approach did not miss relevant studies, systematic reviews were included in the review protocol and search strategy for all review questions. Relevant systematic reviews were used to identify any papers not found through the primary search. Committee members were also consulted to identify studies that were missed. If additional studies were found that were erroneously excluded during the priority screening process, the full database was subsequently screened.

The decision whether or not to use priority screening was taken by the reviewing team depending on the perceived likelihood that stopping criteria would be met, based on the size of the database, heterogeneity of studies included in the review and predicted number of includes. If it was thought that stopping criteria were unlikely to be met, priority screening was not used, and the full database was screened.

The full text of potentially eligible studies was retrieved and assessed according to the criteria specified in the review protocol. A standardised form was used to extract data from included studies.

K.2 Data synthesis for validating prediction models

K.2.1 Pairwise meta-analysis

Pairwise meta-analyses were performed in Cochrane Review Manager V5.3. Where appropriate, hazard ratios were pooled using the generic inverse-variance method. Adjusted odds ratios and hazard ratios from multivariate models were only pooled if

the same set of factors were used across multiple studies and if the same thresholds to measure factors were used across studies. Both odds ratios/hazard ratios and absolute risks were presented, with absolute risks calculated by applying the odds ratio/hazard ratio to the risk in the comparator arm of the meta-analysis (calculated as the total number events in the comparator arms of studies in the meta-analysis divided by the total number of participants in the comparator arms of studies in the meta-analysis).

Random effects models were fitted when there was significant between-study heterogeneity in methodology, population, predictor or comparator was identified by the reviewer in advance of data analysis. This decision was made and recorded before any data analysis was undertaken.

For all other syntheses, fixed- and random-effects models were fitted, with the presented analysis dependent on the degree of heterogeneity in the assembled evidence. Fixed-effects models were the preferred choice to report. Fixed-effects models were deemed to be inappropriate if there was significant statistical heterogeneity in the meta-analysis, defined as $I^2 \geq 50\%$.

In any meta-analyses where some (but not all) of the data came from studies at high risk of bias, a sensitivity analysis was conducted, excluding those studies from the analysis. Results from both the full and restricted meta-analyses are reported. Similarly, in any meta-analyses where some (but not all) of the data came from indirect studies, a sensitivity analysis was conducted, excluding those studies from the analysis.

K.2.2 Appraising the quality of evidence

Studies evaluating prediction models

Individual studies validating prediction models were assessed using the PROBAST checklist. Each individual study was classified into one of the following three groups:

- Low risk of bias – The true effect size for the study is likely to be close to the estimated effect size.
- Moderate risk of bias – There is a possibility the true effect size for the study is substantially different to the estimated effect size.
- High risk of bias – It is likely the true effect size for the study is substantially different to the estimated effect size.

Each individual study was also classified into one of three groups for directness, based on if there were concerns about the population, predictor/comparator/outcome to be predicted in the study and how directly these variables could address the specified review question. Studies were rated as follows:

- Direct – No important deviations from the protocol in population, predictor/comparator/outcome to be predicted.
- Partially indirect – Important deviations from the protocol in one of the population, predictor/comparator/outcome to be predicted.
- Indirect – Important deviations from the protocol in at least two of the population, predictor/comparator/outcome to be predicted.

Modified GRADE for prediction models

GRADE has not been developed for use with data from prediction models, therefore a modified approach was applied using the GRADE framework. The approach taken depended on the outcome data produced by the decision model. Measures of association (such as

HRs or ORs) were assessed as described below in the section on quality assessment of association studies (see Modified GRADE for association data).

Clinical decision thresholds

The committee were asked to define clinical decision thresholds for association outcomes based on the degree of association that was considered clinically important for decision making. In cases where the committee were unable to define a clinical decision threshold by consensus, the line of no effect was used at the clinical decision threshold for the purpose of rating imprecision in GRADE.

Modified GRADE for association data

GRADE has not been developed for use with association studies, therefore a modified approach was applied using the GRADE framework. Data from cohort studies was initially rated as high quality, with the quality of the evidence for each outcome then downgraded or not from this initial point. For the purpose of rating risk of bias and indirectness, single studies were rated in the same way as meta-analysed studies, but with 100% of the weight in analysis contributed by that single study.

Table 9: Rationale for downgrading quality of evidence for association studies

GRADE criteria	Reasons for downgrading quality
Risk of bias	Not serious: If less than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from studies at high risk of bias, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between studies at high and low risk of bias. In addition, unadjusted odds ratio outcomes from univariate analyses were downgraded one level, in addition to any downgrading for risk of bias in individual studies. Adjusted odds ratios from multivariate analyses were not similarly downgraded.
Indirectness	Not serious: If less than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between direct and indirect studies.
Inconsistency	Concerns about inconsistency of effects across studies, occurring when there is unexplained variability in the treatment effect demonstrated across studies (heterogeneity). This was assessed using the I^2 statistic. N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study. Not serious: If the I^2 was less than 33.3%, the outcome was not downgraded. Serious: If the I^2 was between 33.3% and 66.7%, the outcome was downgraded one level. Very serious: If the I^2 was greater than 66.7%, the outcome was downgraded two levels.

GRADE criteria	Reasons for downgrading quality
	Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between studies with the smallest and largest effect sizes.
Imprecision	If a clinical decision threshold other than the line of no effect was defined for the outcome, the outcome was downgraded once if the 95% confidence interval for the effect size crossed one line of the clinical decision threshold, and twice if it crosses both clinical decision thresholds. If the line of no effect was defined as a clinical decision threshold for the outcome, it was downgraded once if the 95% confidence interval for the effect size crossed the line of no effect (i.e. the outcome was not statistically significant), and twice if the sample size of the study was sufficiently small that it is not plausible any realistic effect size could have been detected. Outcomes meeting the criteria for downgrading above were not downgraded if the confidence interval was sufficiently narrow that the upper and lower bounds would correspond to clinically equivalent scenarios.
Publication bias	If the review team became aware of evidence of publication bias (for example, evidence of unpublished trials where there was evidence that the effect estimate differed in published and unpublished data), the outcome was downgraded once. If no evidence of publication bias was found for any outcomes in a review (as was often the case), this domain was excluded from GRADE profiles to improve readability.

The quality of evidence for each outcome was upgraded if either of the following conditions were met:

- Data showing an effect size sufficiently large that it cannot be explained by confounding alone.
- Data where all plausible residual confounding is likely to increase our confidence in the effect estimate.

K.3 Methods for combining c-statistics

C-statistics were assessed using the categories in Table 10 below.

Table 10 Interpretation of c-statistics

Value of c-statistic	Interpretation
c-statistic <0.6	**Poor classification accuracy**
0.6 ≤ c-statistic <0.7	**Adequate classification accuracy**
0.7 ≤ c-statistic <0.8	**Good classification accuracy**
0.8 ≤ c-statistic <0.9	**Excellent classification accuracy**
0.9 ≤ c-statistic < 1.0	**Outstanding classification accuracy**

Meta-analyses were carried out using the metamisc package in R v4.1.0, which confines the analysis results to between 0 and 1 matching the limited range of values that c-statistics can take. Random effects meta-analysis was used when the I^2 was 50% or greater.

In any meta-analyses where some (but not all) of the data came from studies at high risk of bias, a sensitivity analysis was conducted, excluding those studies from the analysis. Results from both the full and restricted meta-analyses are reported. Similarly, in any meta-analyses where some (but not all) of the data came from indirect studies, a sensitivity analysis was conducted, excluding those studies from the analysis.

K.3.1 Modified GRADE for c-statistics

A modified version of GRADE was carried out to assess the quality of the meta-analysed c-statistics as follows. For the purpose of rating risk of bias and indirectness, single studies were rated in the same way as meta-analysed studies, but with 100% of the weight in analysis contributed by that single study.

Risk of bias

- Not serious: If less than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the overall outcome was not downgraded.
- Serious: If greater than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the outcome was downgraded one level.
- Very serious: If greater than 33.3% of the weight in a meta-analysis came from studies at high risk of bias, the outcome was downgraded two levels.

Indirectness

- Not serious: If less than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the overall outcome was not downgraded.
- Serious: If greater than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the outcome was downgraded one level.
- Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels.

Inconsistency

Concerns about inconsistency of effects across studies, occurring when there is unexplained variability in the treatment effect demonstrated across studies (heterogeneity), after appropriate pre-specified subgroup analyses have been conducted. This was assessed using the I^2 statistic.

- N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study.
- Not serious: If the I^2 was less than 33.3%, the outcome was not downgraded.
- Serious: If the I^2 was between 33.3% and 66.7%, the outcome was downgraded one level.
- Very serious: If the I^2 was greater than 66.7%, the outcome was downgraded two levels.

Imprecision

The 95% CI boundaries were examined and if they crossed 2 categories of test classification accuracy then the study was downgraded once (imprecision rated as serious); if the boundaries crossed 3 categories then the study was downgraded twice (very serious imprecision).

In cases where meta-analyses could not be carried out due to single studies with or without 95% CI, the following decision rules were used to assess risk of bias, indirectness, imprecision and inconsistency for each outcome:

1. Risk of bias and indirectness were assessed as detailed above.
2. Imprecision
    - Single study with 95% CI: the 95% CI boundaries were examined and if they crossed 2 categories of test classification accuracy then the study was

downgraded once (imprecision rated as serious); if the boundaries crossed 3 categories then the study was downgraded twice (very serious imprecision).
    - Single study without 95% CI: the mean sample size was calculated and if this was < 250 then the analysis was downgraded twice (very serious); if it was >250, but > 500 the analysis was downgraded once (serious); if the mean was > 500 people/study then the analysis was not downgraded (not serious).
3. Inconsistency
    - Single study with or without 95% CI: N/A.

K.3.2 Methods for combining Brier scores

Brier scores were considered separately for each study and not combined in a meta-analysis.

K.3.3 Modified GRADE for Brier scores

Risk of bias

- Not serious: If the study was at low risk of bias, the outcome was not downgraded.
- Serious: If the study was at moderate risk of bias the outcomes was downgraded one level.
- Very serious: If the study was at high risk of bias, the outcome was downgraded two levels.

Indirectness

- Not serious: If the study was directly applicable, the overall outcome was not downgraded.
- Serious: If the study was partially indirect the outcome was downgraded one level.
- Very serious: If the study was indirect, the outcome was downgraded two levels.

Inconsistency

- N/A: studies were not pooled.

Imprecision

The 95% CI boundaries were examined and imprecision was downgraded one level if the extent of the confidence intervals had a serious impact on the certainty of the committee in the effect estimate for decision making. Imprecision was downgraded 2 levels if the extent of the confidence intervals had a very serious impact on the certainty of the committee in the effect estimate for decision making

Appendix L – Prostate cancer risk stratification models

Table 11: Prostate cancer risk stratification models with criteria to categorise risk

Risk stratification model	Tiers					References
3 tier prostate cancer risk stratification models						
NICE		**Low risk** PSA <10 ng/ml and GS ≤6 and cT1 to T2a	**Intermediate risk** PSA 10 to 20 ng/ml or GS 7 or cT2b	**High risk** PSA >20 ng/ml or GS 8 to 10 or ≥cT2c		NICE NG131
D'Amico		**Low risk** PSA <10 ng/ml and GS ≤6 and cT1c-T2a	**Intermediate risk** PSA 10 to 20 ng/ml or GS 7 or cT2b	**High risk** PSA >20 ng/ml or GS 8 to 10 or cT2c		Zelic 2020
EUA		**Low** PSA <10 ng/ml and GS ≤6 (ISUP 1) and cT1c-T2a	**Intermediate** PSA 10 to 20 ng/ml or GS 7 (ISUP 2 to 3) or cT2b	**High** PSA >20 ng/ml or GS >7 (ISUP 4 to 5) or cT2c		Zelic 2020
GUROC		**Low** PSA ≤10 ng/ml and GS ≤6 and cT1-T2a	**Intermediate** PSA ≤20 ng/ml and GS ≤7 and cT1-T2 not otherwise low risk	**High** PSA >20 ng/ml or GS 8 to 10 or ≥cT3a		Zelic 2020
5 tier prostate cancer risk stratification models						
CPG	**CPG1** GS 6 (ISUP 1) and PSA <10 ng/ml and cT1-T2	**CPG2** GS 3+4=7 (ISUP 2) or PSA 10 to 20 ng/ml and	**CPG3** GS 3+4=7 (ISUP 2) and PSA 10 to 20 ng/ml and	**CPG4** GS 8 (ISUP 4) or PSA >20 ng/ml or cT3	**CPG5** Any combination of GS 8 (ISUP 4), PSA >20 ng/ml or cT3	Gnanapragasam 2018

Risk stratification model	Tiers					References
		cT1-T2	cT1-T2 **OR** GS 4+3=7 (ISUP 3) and cT1-T2		**OR** GS 9 to 10 (ISUP 5) or cT4	
AUA-i	**Very low** PSA <10 ng/ml and ISUP 1 and cT1-T2a and <34% positive cores and no cores with >50% cancer and PSAD <0.15	**Low** PSA <10 ng/ml and ISUP 1 and cT1-T2a	**Favourable intermediate** ISUP 1 and PSA 10 to <20 ng/ml **OR** ISUP 2 and PSA <10 ng/ml	**Unfavourable intermediate** ISUP 2 and PSA 10 to <20 ng/ml or cT2b-T2c **OR** ISUP 3 and PSA <20 ng/ml	**High** PSA >20 ng/ml or ISUP 4 to 5 or ≥cT3	Zelic 2020
Modified D'Amico (incorporation of percent of positive cores)	**Low risk** Same as the traditional D'Amico group	**Intermediate risk 1** CPR ≤50%	**Intermediate risk 2** CPR >50%	**High risk 1** CPR ≤50%	**High risk 2** CPR >50%	Abdel-Rahman 2018

American Urological Association with favourable and non-favourable intermediate groups (AUA-i), Cambridge Prognostic Groups (CPG), core positive ratio (CPR), cT (clinical stage), European Association of Urology (EAU), Genitourinary Radiation Oncologists of Canada (GUROC), GS (Gleason score), ISUP (International Society of Urological Pathology grade group), NICE (National Institute for Health and Care Excellence), PSA (prostate-specific antigen), PSAD (prostate-specific antigen density)

Appendix M - Research Recommendation

M.1.1 Research recommendation

What is the diagnostic accuracy of staging investigations in people with CPG 2 and 3 prostate cancer?

M.1.2 Why this is important

The committee considered how recommendations on bone scans were impacted by the CPG stratification scheme that is now recommended. They highlighted the lack of evidence for staging investigations for people with CPG 2 and 3 localised prostate cancer. Research in this area will inform future updates of the guideline.

M.1.3 Rationale for research recommendation

Importance to 'patients' or the population	Having the right staging investigations will give patients a more accurate prognosis and will allow treatments to be correctly targeted, minimising both over and under treatment.
Relevance to NICE guidance	The NICE guideline on prostate cancer does not currently provide recommendations on staging investigations for CPG 2 and 3 prostate cancer. The research recommendation would inform future guideline updates in this area.
Relevance to the NHS	Access to the correct staging investigations will allow treatments to be correctly targeted.
National priorities	High
Current evidence base	Evidence is available on the use of staging investigations to detect metastases for people with high risk disease, but the use of staging investigations for CPG 2 and 3 prostate cancer is uncertain.
Equality considerations	No specific equalities considerations were identified for this research recommendation.

M.1.4 Modified PICO table

Population	People with CPG 2 and 3 localised prostate cancer
Index test	Staging investigations (for example, bone scans, PSMA scans)
Reference Standard	Histological confirmation of metastatic disease
Outcomes	Diagnostic test accuracy outcomes including sensitivity, specificity and likelihood ratios Proportion of patients with change in management due to outcome of staging investigation
Study design	Cross sectional
Timeframe	Not applicable
Stratification	Stratification by CPG group