Australian Government

National Health and Medical Research Council

Prostate-Specific Antigen (PSA) testing in asymptomatic men: Technical Report 2013

© Commonwealth of Australia 2013

Electronic document

This work is copyright. You may download, display, print and reproduce the whole or part of this work in unaltered form for your own personal use or, if you are part of an organisation, for internal use within your organisation, but only if you or your organisation do not use the reproduction for any commercial purpose and retain this copyright notice and all disclaimer notices as part of that reproduction. Apart from rights to use as permitted by the *Copyright Act 1968* or allowed by this copyright notice, all other rights are reserved and you are not allowed to reproduce the whole or any part of this work in any way (electronic or otherwise) without first being given the specific written permission from the Commonwealth to do so. Requests and enquiries concerning reproduction and rights are to be sent to Strategic Communications, National Health and Medical Research Council, GPO Box 1421, Canberra ACT 2600 or via email to nhmrc.publications@nhmrc.gov.au.

ISBN Online: 1864965991

Suggested citation

National Health and Medical Research Council (2013). *Prostate-Specific Antigen (PSA) testing in asymptomatic men: Technical Report.* Canberra: National Health and Medical Research Council.

Disclaimer

This report should be read in conjunction with *Prostate-Specific Antigen (PSA) testing in asymptomatic men: Evidence Evaluation Report.*

Contact:

National Health and Medical Research Council
16 Marcus Clarke Street
Canberra ACT 2601

GPO Box 1421
Canberra ACT 2601

Phone: 61 2 6217 9000
Fax: 61 2 6217 9100
Email: nhmrc@nhmrc.gov.au
Web: www.nhmrc.gov.au

Available from: www.nhmrc.gov.au/guidelines/publications/men4

Published: July 2013

NHMRC Reference code: MEN4b

Acknowledgements

This report was co-funded by the Australian Government Department of Health and Ageing (DoHA) and the National Health and Medical Research Council (NHMRC). It was prepared by Optum (the evidence reviewer), in conjunction with the NHMRC PSA Testing Expert Advisory Group (EAG).

The Optum Project Team included:
Dr Suzanne Campbell, Associate Director
Dr Jennifer Wong, Research Analyst
Ms Jyoti Rauniyar, Research Analyst

Members of the PSA Testing Expert Advisory Group were:

Professor Adele Green (Chair)	Senior Scientist, Queensland Institute of Medical Research
Professor Bruce Armstrong	School of Public Health, University of Sydney
Professor Dallas English	Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne
Professor Ian Olver	Chief Executive Officer, Cancer Council Australia
Professor Judith Clements	Institute of Health and Biomedical Innovation, Queensland University of Technology
Emeritus Professor Villis Marshall	General Manager , Royal Adelaide Hospital
Professor Robert Gardiner	UQ Centre for Clinical Research, Royal Brisbane and Women's Hospital
Professor Jon Emery	School of Primary, Aboriginal and Rural Health Care, University of Western Australia
Dr Helen Zorbas	Chief Executive Officer, Cancer Australia
Professor Paul Glasziou	Centre for Research in Evidence-Based Practice, Bond University
Professor James Kench	Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital
Mr David Sandoe	Prostate Cancer Foundation of Australia
Mr John Stubbs	Canspeak, Sydney
Professor Madeline King	School of Psychology, the University of Sydney

The NHMRC project team included:
Ms Cathy Connor
Ms Melissa Lawrance
Ms Emma Breen

The DoHA project team included:
Ms Toni Smith
Mr Chris Milton
Ms Kim Dobbie
Dr Tracey Bessell

An independent methodological eview was conducted by Professor Jenny Doust, Centre for Research in Evidence-Based Practice, Bond University.

Table of Contents

List of Tables

List of Abbreviations

AHRQ	Agency for Healthcare Research and Quality
CADTH	Canadian Agency for Drugs and Technologies in Health
CEO	Chief Executive Officer
DRE	Digital rectal examination
EAG	Expert Advisory Group
EMBASE	Excerpta Medica Database
ERSPC	European Randomised Study of Screening for Prostate Cancer
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HCC	Health Care Committee
HTA	Health technology assessment
NHMRC	National Health and Medical Research Council
NICE	National Institute for Health and Clinical Excellence
NZGG	New Zealand Guidelines Group
PCHC	Prevention and Community Health Committee
PICO	Population, Intervention, Comparator, Outcome
PSA	Prostate-specific antigen
RCT	Randomised controlled trial
USPTF	The United States Preventive Services Task Force

Introduction

This Technical Report accompanies the National Health and Medical Research Council (NHMRC) Evidence Evaluation Report on prostate-specific antigen (PSA) testing in asymptomatic men. It has been prepared by Optum (the evidence reviewer, formerly Health Technology Analysts Pty Ltd), in conjunction with the PSA Testing Expert Advisory Group (EAG).

Section 1 of the Technical Report provides a comprehensive description of the methods that were used to systematically review the systematic reviews (Level I evidence) that assessed the effectiveness of using the PSA test for reducing mortality and morbidity due to prostate cancer in asymptomatic men. Specifically, it provides:

- The primary clinical research questions that were used to define the systematic evidence review
- The literature search strategies that were used to identify studies relevant to the primary clinical research questions
- A description of the methodology that was used to critically appraise the evidence relevant to the primary clinical research questions
- Quality assessment and data extraction forms for studies relevant to the primary clinical research questions
- Evidence Statement Forms for each outcome of the primary clinical research questions
- A brief description of how comments from the EAG, NHMRC's relevant Principal Committees, Council of NHMRC and independent expert review have been addressed

Section 2 of the Technical Report provides a description of the research questions and methodology that were used for the supplementary non-systematic literature review.

Aboriginal and Torres Strait Islander populations

The review of evidence relating to PSA testing in asymptomatic men did not specifically search for, or limit, the retrieval of articles to studies that included the Aboriginal or Torres Strait Islander peoples. However, the evidence reviewers did identify any papers that addressed these populations for specific consideration by the EAG.

The evidence reviewer notes that no relevant socioeconomic literature pertaining to Australia's Indigenous population was identified in the literature searches for either the systematic or non-systematic evidence reviews.

1 Methodology for systematic review

For this evaluation, the NHMRC defined a systematic review as 'a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies. (*http://www.cochrane.org/glossary/5*). Systematic reviews should aim to identify all studies addressing the question, regardless as to whether or not it has been published. As a minimum unpublished literature should include trials registered on clinical trial databases'.

1.1 Research question development

The clinical research questions to be addressed by the systematic review were developed and agreed upon by the EAG at a meeting held on 24 August 2012. These primary clinical research questions were structured and scoped according to the PICO criteria (Population, Intervention, Comparator, Outcome). Use of the PICO framework facilitates the systematic review process as it improves conceptual clarity of the clinical problem, allows more complex search strategies, results in more precise search results, and allows evidence to be selected appropriately.

The two primary clinical research questions that were developed for the systematic review are:

1. Does PSA testing, with or without digital rectal examination (DRE), in asymptomatic men reduce prostate cancer-specific mortality or all-cause mortality?
2. Does PSA testing, with or without DRE, in asymptomatic men reduce morbidity due to advanced prostate cancer?

The agreed PICO criteria for the primary clinical research questions are detailed in **Table 1**.

Table 1 PICO criteria for the primary clinical research questions developed for the systematic review

Primary clinical research question 1: ***Does PSA testing, with or without digital rectal examination, in asymptomatic men reduce prostate cancer-specific mortality or all-cause mortality?***				
Population	**Intervention**	**Comparison**	**Outcomes**	**Other systematic review considerations**
Asymptomatic men Stratify by: • Age (however reported)[a] • +/- risk factors (e.g. older age, family history of prostate cancer) • Comorbidities (including life expectancy)	PSA testing (all modalities) with or without DRE	No PSA testing (control)[b]	• Mortality (prostate cancer-specific, all-cause), relative to time since testing	Limits: • Search period: 2002-04 Sept 2012 • Restrict to Level I and II evidence • Full length publications only • English only publications[c]

Primary clinical research question 2:
Does PSA testing, with or without digital rectal examination, in asymptomatic men reduce morbidity due to advanced prostate cancer?

Population	Intervention	Comparison	Outcomes	Other systematic review considerations
Asymptomatic men Stratify by: • Age (however reported)[a] • +/- risk factors (e.g. older age, family history of prostate cancer) • Comorbidities (including life expectancy)	PSA testing (all modalities) with or without DRE	No PSA testing (control)[b]	• Prostate cancer-specific metastatic disease • Skeletal-related events (e.g. osteoporosis, fractures) • QoL[d]	Limits: • Search period: 2002-04 Sept 2012 • Restrict to Level I and II evidence • Full length publications only • English only publications[c]

Abbreviations: DRE, digital rectal examination; PSA, prostate-specific antigen; QoL, quality of life

[a] The appropriate cut-off in Australia will be discussed and agreed by the EAG when considering the evidence. Presentation of results will be limited by the eligibility criteria of the studies and how they report their results.

[b] Includes DRE alone if the intervention is PSA testing + DRE.

[c] Studies will be excluded if they are not published in the English language. However, non-English publications that otherwise fulfil the eligibility criteria will be brought to the attention of the EAG.

[d] The QoL outcome will encompass detailed data extracted from PRO instruments (e.g. SF-36, UCLA-PCI, QLQ-C30), including instruments for assessment of anxiety (e.g. STAI). Subdomain data will be extracted. Rates of sexual dysfunction, bowel function, urinary incontinence, etc. will not be extracted (except for the purposes of the non-systematic literature review).

1.2 Literature searches

Systematic literature searches were conducted for each of the primary clinical research questions in accordance with the NHMRC standards for evidence review. The literature search strategies were developed based on the NHMRC levels of evidence hierarchy (**Table 2**). Level I evidence refers to systematic reviews of Level II evidence, and is considered to be the highest level of evidence for intervention and screening intervention questions. Level II evidence refers to randomised controlled trials (RCTs). RCTs allow the comparison of groups under investigation whilst attempting to minimise bias. They are considered to be the only reliable type of clinical trial design for evaluating the most important outcome for PSA testing, prostate cancer-specific mortality.

To ensure that the systematic review included the most recent evidence, the literature search was performed in two stages:

1. A systematic review of Level I evidence relating to each of the primary clinical research questions.
2. A literature search update designed to identify Level II studies that were published after the search date of the most comprehensive and highest quality Level I evidence available. This pivotal review was determined to be Ilic et al (2010), which was later superseded by Ilic et al (2013), hereafter known as the Cochrane review.

Table 2 NHMRC levels of evidence hierarchy

Level	Intervention[b]	Screening intervention[b]
I[a]	A systematic review of Level II studies	A systematic review of Level II studies
II	A randomised controlled trial	A randomised controlled trial
III-1	A pseudo-randomised controlled trial (i.e., alternate allocation or some other method)	A pseudo-randomised controlled trial (i.e., alternate allocation or some other method)
III-2	A comparative study with concurrent controls: • Non-randomised, experimental trial[c] • Cohort study • Case-control study • Interrupted time series with a control group	A comparative study with concurrent controls: • Non-randomised, experimental trial[c] • Cohort study • Case-control study
III-3	A comparative study without concurrent controls: • Historical control study • Two or more single arm study[d] • Interrupted time series without a parallel control group	A comparative study without concurrent controls: • Historical control study • Two or more single arm study[d]
IV	Case series with either post-test or pre-test/post-test outcomes	Case series

Source: NHMRC (2009)

[a] A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of Level II evidence. Systematic reviews of Level II evidence provide more data than the individual studies and any meta-analyses will increase the precision of the overall results, reducing the likelihood that the results are affected by chance. Systematic reviews of lower level evidence present results of likely poor internal validity and thus are rated on the likelihood that the results have been affected by bias, rather than whether the systematic review itself is of good quality. Systematic review quality should be assessed separately. A systematic review should consist of at least two studies. In systematic reviews that include different study designs, the overall level of evidence should relate to each individual outcome/result, as different studies (and study designs) might contribute to each different outcome.

[b] Definitions of these study designs are provided on pages 7–8 of the NHMRC toolkit, *How to use the evidence: assessment and application of scientific evidence* (NHMRC 2000).

[c] This also includes controlled before-and-after (pre-test/post-test) studies, as well as indirect comparisons (i.e. utilise A vs. B and B vs. C, to determine A vs. C).

[d] Comparing single arm studies i.e., case series from two studies. This would also include unadjusted indirect comparisons (i.e. utilise A vs. B and B vs. C, to determine A vs. C but where there is no statistical adjustment for B).

The literature search strategies used to identify publications relevant to the primary clinical research questions are shown in **Table 3** and **Table 4**. The literature searches were primarily conducted using EMBASE.com (which searches EMBASE and Medline databases concurrently), the Cochrane Library, and PubMed. Additional health technology assessment (HTA) agency websites (e.g. National Institute for Health and Clinical Excellence [NICE] in the United Kingdom, Canadian Agency for Drugs and Technologies in Health [CADTH]), and guideline databases (e.g. Guidelines International Network [GIN], National Guidelines Clearing House) were also searched. After reviewing the initially retrieved citations, a manual search of the reference lists of relevant papers was also performed.

Table 3 Literature search strategy for the systematic review of Level I evidence

Database	#	Search terms	Citations retrieved	Total number of citations
Medline and EMBASE (using EMBASE.com interface)[a] Search date: 4 September 2012	#1	'prostate specific antigen' OR 'prostate specific antigen'/exp	33,511	592
	#2	psa:ab,ti	29,631	
	#3	#1 OR #2 NOT 'psoriatic arthritis' NOT 'psa-ncam'	43,605	
	#4	'meta analysis'/exp OR 'meta analysis' OR 'systematic review'/exp OR 'systematic review' OR 'pooled analysis' OR ('review'/exp OR 'review' AND (systemat* OR pool*))	180,149	
	#5	#3 AND #4	706	
	#6	#5 AND [2002-2013]/py	592	
Cochrane Library (systematic reviews, other reviews and health technology assessments) Search date: 4 September 2012	#1	"prostate specific antigen" in Title, Abstract or Keywords, from 2002 to 2013 in Cochrane Database of Systematic Reviews	10	40
	#2	"prostate specific antigen" in Title, Abstract or Keywords, from 2002 to 2013 in Database of Abstracts of Reviews of Effects	20	
	#3	"prostate specific antigen" in Title, Abstract or Keywords, from 2002 to 2013 in Health Technology Assessment Database	10	
Citations identified through other HTA agency and guideline websites • NICE • CADTH • AHRQ • GIN • National Guidelines Clearinghouse			3	3
Citations identified through manual check of reference lists			0	0
Total number of citations identified				**635**

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; CADTH, Canadian Agency for Drugs and Technologies in Health; GIN, Guidelines International Network; NICE, National Institute for Health and Clinical Excellence

[a] Records are included in EMBASE.com as soon as the citation and abstract is available from the publisher. Although the full indexing is not yet available, In-Process records are enriched with index terms automatically generated from title and abstract. In some cases In-Process records themselves replace Articles in Press.

Table 4 Literature search strategy for the Level II evidence update

Database	#	Search terms[a]	Citations retrieved	Total number of citations
Medline and EMBASE (using EMBASE.com interface)[b] Search date: 4 September 2012	#1	'prostate specific antigen' OR 'prostate specific antigen'/exp	33,511	1601
	#2	psa:ab,ti	29,631	
	#3	#1 OR #2 NOT 'psoriatic arthritis' NOT 'psa-ncam'	43,605	
	#4	'comparative study'/exp OR 'comparative study' OR 'clinical trial'/exp OR 'clinical trial' OR 'randomized controlled trial'/exp OR 'randomization'/exp OR 'single blind procedure'/exp OR 'single blind procedure' OR 'double blind procedure'/exp OR 'double blind procedure' OR 'triple blind procedure'/exp OR 'triple blind procedure' OR 'crossover procedure'/exp OR 'crossover procedure' OR 'placebo'/exp OR placebo* OR random* OR rct OR 'single blind' OR 'single blinded' OR 'double blind' OR 'double blinded' OR 'treble blind' OR 'treble blinded' OR 'triple blind' OR 'triple blinded' OR 'prospective study'/exp OR 'prospective study'	2,606,900	
	#5	#3 AND #4	11,610	
	#6	#5 AND [1-6-2010]/sd	3231	
	#7	#6 AND ('mass screening'/exp OR screening OR asymptomatic OR healthy OR 'screening test'/exp OR test*)	1601	
Cochrane Library (clinical trials) Search date: 4 September 2012	#1	"prostate specific antigen" in Title, Abstract or Keywords, from 1 June 2010 in Cochrane Central Register of Controlled Trials	83	83
PubMed (for 'Epub ahead of print' advanced publications 2012-2013) Search date: 4 September 2012	#1	(prostate specific antigen[Title/Abstract]) OR (PSA[Title/Abstract])	26,137	101
	#2	#1 AND publication date from 1 July 2012	586	
	#3	#2 AND (comparative study OR clinical trial OR randomized controlled trial OR randomization OR crossover OR placebo* OR random* OR rct OR single blind OR single blinded OR double blind OR double blinded OR treble blind OR treble blinded OR triple blind OR triple blinded OR prospective study))	120	
	#4	#3 NOT psoriatic arthritis NOT PSA-NCAM	101	
Citations identified through manual check of reference lists			0	0
Total number of citations identified				**1785**

[a] This search strategy was initially used to update the literature search performed by Ilic et al in July 2010 for the Cochrane Collaboration review of screening strategies for prostate cancer (Ilic et al, 2010).

[b] Records are included in EMBASE.com as soon as the citation and abstract is available from the publisher. Although the full indexing is not yet available, In-Process records are enriched with index terms automatically generated from title and abstract. In some cases In-Process records themselves replace Articles in Press.

1.3 Study eligibility

All citations identified in the literature searches described in **Table 3** and **Table 4** were reviewed based on information in the publication title and, where available, the abstract. Relevant publications

were retrieved and reviewed in full text before a final decision was made on their inclusion or exclusion for the systematic review.

Consistent with the PICO criteria for the systematic review (**Table 1**), the following *a priori* exclusion criteria were applied to specifically identify relevant systematic reviews or RCTs:

- Duplicate citation
- Wrong study type: the article was not a systematic review or RCT (e.g. a narrative review, editorial, letter, case report, conference abstract, observational or non-comparative study)
- Wrong population: the study was not in asymptomatic men
- Wrong intervention: the intervention was not PSA testing (of any modality) with or without DRE
- Superseded by a more recent systematic review by the same organisation/authors
- Not in English

The application of the exclusion criteria to citations identified through the systematic literature searches is shown in **Table 5.** Studies that were excluded after full text review are documented with their reasons for exclusion in **Appendix A**. Examples of reasons for exclusion in this circumstance include studies that did not include outcomes relevant to the primary research questions, and those that did not include adequate methodology details.

One Level I study was excluded prior to full text review as it was not published in English. The citation details for this publication are:

> Bastos Varzim CA, Srulzon GB, Macedo Cortado PL, Rodrigues N, Jr. (2004) Importance of digital rectal examination and PSA in early prostate cancer diagnosis. *Revista Brasileira de Medicina* 61(7):471-4.

Table 5 Summary of citations retrieved in the systematic review of Level I evidence

Literature search for Level I evidence	**Total number of citations**
Total number of citations identified	635
Citations excluded after title/abstract review: • Duplicate citation • Wrong study type • Wrong population • Wrong intervention • Superseded by a more recent systematic review by the same HTA agency/authors • Not in English	 40 398 73 100 5 1
Number of studies reviewed in full text	18
Studies excluded after full text review[a]	11
Final number of eligible Level I studies	**7**
Literature search for Level II evidence (published after the key systematic review)	**Total number of citations**
Total number of citations identified	1785
Citations excluded after title/abstract review: • Duplicate citation • Wrong study type • Wrong population • Wrong intervention	 100 1184 386 87
Number of studies reviewed in full text	28
Studies excluded after full text review[a]	16
Final number of citations referring to Level II studies	12

[a] Studies excluded after full text review are documented, with their reasons for exclusion in **Appendix A**.

The literature search identified seven eligible Level I studies (**Appendix B**), including a 2010 Cochrane review of PSA screening for prostate cancer (Ilic et al, 2010). This Cochrane review was known by the EAG and NHMRC to be in the process of being updated, with the updated version due for release in late September 2012. At a meeting held on 24 August 2012, the EAG and NHMRC agreed that the update of the Cochrane review was critical to the evidence review of PSA testing in asymptomatic men. Accordingly, a decision was made to incorporate the updated Cochrane review in the systematic review of Level I evidence. Importantly, the Cochrane literature search was conducted prior to the literature search for the current evidence review. The updated Cochrane review (Ilic et al, 2013) became available to the NHMRC in draft form on 22 November 2012 and was published on 31 January 2013. It supersedes the 2010 version initially identified in the Cochrane Library database.

Although the literature search for Level II evidence published after the key systematic review did not identify any new RCTs, 12 follow-up publications were identified for the RCTs already identified in the systematic reviews. These publications are included in the list of included citations (**Appendix B**).

1.4 Critical appraisal

Studies identified for inclusion from the literature search were classified according to the NHMRC dimensions of evidence (**Table 6**). There are three main domains within the NHMRC dimensions of

evidence: strength of the evidence, size of the effect, and relevance of the evidence. The strength of the evidence was derived directly from the literature identified for a particular interventional study. Assessment of the size of the effect and the relevance of the evidence was discussed with the EAG at a meeting held on 22 January 2013.

Study quality was evaluated and reported based on the quality assessment questions included in the NHMRC toolkit, *How to use the evidence: assessment and application of scientific evidence* (NHMRC 2009). The quality assessment forms for the included Level I studies are presented in **Section 1.7.1**. The quality assessment forms for the RCTs that are included within the Level I studies are presented in **Section 1.7.2**.

Table 6 NHMRC dimensions of evidence

Dimension	Definition
Strength of evidence	
• Level[a]	Each included study is assessed according to its place in the research hierarchy. This illustrates the potential of each included study to adequately answer a particular research question and indicates the degree to which design has minimised the impact of bias on the results.
• Quality[b]	Included studies are critically appraised for methodological quality. Each study is assessed according to the potential that bias, confounding and/or chance has influenced the results.
• Statistical precision	Primary outcomes of included studies are assessed to establish if the effect is real, rather than due to chance. Using a level of significance such as a P-value and/or confidence interval the precision of the estimate of the effect is evaluated. This considers the degree of certainty regarding the existence of a true effect.
Size of effect	The clinical importance of the findings of each study is assessed. This concept refers to the measure of effect or point estimate reported in the results of each study (e.g. mean difference, relative risk etc). For meta-analysis pooled measures of effect are assessed. Size of effect refers to the distance of the point estimate from its null value and also the values included in the corresponding 95% confidence interval. Size of effect indicates the clinical impact a particular factor or intervention will have on a patient and is considered in the context of patient relevant clinical differences.
Relevance of evidence	The translation of research evidence to clinical practice is addressed by this dimension. It is regarded as potentially the most subjective of the evidence assessments. There are two questions concerning the appropriateness of outcomes and relevance of study questions: • Are the outcomes measured in the study relevant to patients? • How closely do the elements of the study research question match with those of the clinical question being considered?

Source: NHMRC (2009)

[a] The level of evidence for each study was determined using the NHMRC levels of evidence hierarchy outlined in Table 2.

[b] Study quality was evaluated and reported based on the quality assessment questions which are included in the NHMRC toolkit, How to use the evidence: assessment and application of scientific evidence (NHMRC 2000).

1.5 Data extraction

Standardised data extraction forms and evidence summary tables were used to capture information relevant to the systematic review of PSA testing in accordance with NHMRC standards. Extracted information included:

- General study details (citation, study design, evidence level, country and setting)
- Affiliations/sources of funds for each of the included studies
- Internal and external validity considerations
- Details of the PSA test (including PSA cut offs)
- Participant details, including key demographic characteristics
- Primary, secondary and other study outcome results

The data were extracted by one evidence reviewer, with the completed data extraction forms checked by a second, independent evidence reviewer. Data extraction was only completed for Level I studies (systematic reviews of Level II evidence) that were assessed as good or fair quality by the evidence reviewer and not for poor quality studies. Level I studies are notorious for lacking details of individual studies contained within. In order to capture the information specified in the data extraction forms and assist the EAG with interpretation of the evidence, further details of the Level II

studies were sought from the primary sources. Data extraction forms for all of the included studies are presented in **Section 1.8**. Sources of funding and declared interests of the authors in each included Level I study are tabulated in **Appendix C**.

1.6 Assessment of the body of evidence

As the purpose of this evaluation is for an information paper rather than the development of clinical practice guidelines, modified NHMRC Evidence Statement Forms were used to consolidate the body of evidence relating to each primary clinical research question. Separate Evidence Statement Forms were prepared for each outcome specified in the PICO criteria for each research question. Completed Evidence Statement Forms are presented in Section 1.9.

The NHMRC Evidence Statement Form has five key components: evidence base, consistency, clinical impact, generalisability and applicability (Table 7). Each component was rated according to the body of evidence matrix shown in Table 8. The first two components of the Evidence Statement Form – evidence base and consistency – consider the internal validity of included studies. These components were rated by the evidence reviewer directly based on the available literature. At a meeting held on 22 January 2013, it was agreed by the NHMRC and EAG that the clinical impact, generalisability and applicability components would not require rating for the purposes of this evaluation. It was also decided that an Evidence Statement Form (and consequently, an evidence statement) for the skeletal-related events outcome was not required. Otherwise, the evidence relating to each outcome for each clinical research question was synthesised into one or more evidence statements.

Table 7 Components of the NHMRC Evidence Statement Form

Dimension	Definition
Evidence base	
• Quantity	Reflects the number of studies included as the evidence base. Also takes into account the number of patients in relation to frequency of the outcomes measures (i.e. study statistical power). Meta-analysis can be used to combine results of studies to increase the power and statistical precision of effect estimates.
• Level	Reflects the best study type for the specific type of research question (intervention, prognosis). Level I evidence would be the best evidence to answer each question.
• Quality	Reflects how well studies were designed and conducted in order to eliminate bias.
Consistency	Assesses whether findings are consistent across included studies, including a range of study populations and study designs. Meta-analysis of randomised studies should present statistical analysis of heterogeneity that demonstrates little statistical difference between studies. Presentation of an I^2 statistic illustrates the extent of heterogeneity between studies[a]. Clinical heterogeneity between studies should also be explored.
Clinical impact	Measures the potential benefit from application of the guideline to a population. Several factors need to be considered when estimating clinical impact. These include: relevance of the evidence to the clinical question; statistical precision and size of the effect; relevance of the effect to patients compared with other management options or none. Other relevant factors are the duration of therapy required to achieve the effect and the balance of risks and benefits (taking into account the size of the patient population).

Dimension	Definition
Generalisability	The translation of research evidence to clinical practice is addressed by this dimension. It is regarded as potentially the most subjective of the evidence assessments. There are two questions concerning the appropriateness of outcomes and relevance of study questions: • Are the outcomes measured in the study relevant to patients? • How closely do the elements of the study research question match with those of the clinical question being considered?
Applicability	Addresses whether the evidence base is relevant to the Australian healthcare setting in general or to more local settings for specific recommendations (e.g. rural areas or cities). Factors that will impact the applicability of study findings include organisational factors (e.g. availability of trained staff, specialised equipment and resources) and cultural factors (e.g. attitudes to health issues, including those that may affect compliance with guideline recommendations).

Source: NHMRC (2009)

[a] Most statistical tests of heterogeneity assess whether heterogeneity exists between studies, in contrast I2 quantifies how much heterogeneity exists between studies.

[b] Study quality was evaluated and reported based on the quality assessment questions which are included in the NHMRC toolkit, How to use the evidence: assessment and application of scientific evidence (NHMRC 2000).

Table 8 Body of evidence grading matrix

Component	A	B	C	D
	Excellent	**Good**	**Satisfactory**	**Poor**
Evidence base[a]	One or more Level I studies with a low risk of bias or several Level II studies with a low risk of bias	One or two Level II studies with a low risk of bias or a systematic review/several Level III studies with a low risk of bias	One or two Level III studies with low risk of bias, or Level I or II studies with a moderate risk of bias	Level IV studies, or Level I to III studies with a high risk of bias
Consistency[b]	All studies consistent	Most studies consistent and inconsistency may be explained	Some inconsistency reflecting genuine uncertainty around clinical question	Evidence is inconsistent
Clinical impact[c]	Very large	Substantial	Moderate	Slight/ restricted
Generalisability	Population/s studied in body of evidence are the same as the target population for the guideline	Population/s studied in the body of evidence are similar to the target population for the guideline	Population/s studied in body of evidence different to target population but it is clinically sensible to apply this evidence to target population	Population/s studied in body of evidence different to target population and hard to judge whether it is sensible to generalise to target population
Applicability	Directly applicable to Australian healthcare context	Applicable to Australian healthcare context with few caveats	Probably applicable to Australian healthcare context with some caveats	Not applicable to Australian healthcare context

Source: NHMRC (2009)

[a] Level of evidence determined from the NHMRC evidence hierarchy (**Table 2**)

[b] If there is only one study, this component is ranked as 'not applicable'.

[c] An additional not applicable (NA) category is used when the evidence shows no difference, is conflicting, or underpowered. This is a modification of currently used NHMRC body of evidence matrix.

1.7 Quality assessment forms

1.7.1 Systematic reviews

Y	N	NR	NA	Quality criteria	Error category[a]
				Study ID	**Basch et al (2012)**
				Citation	Basch E, Oliver TK, Vickers A, Thompson I, Kantoff P, Parnes H, Loblaw DA, Roth B, Williams J, Nam RK (2012) Screening for Prostate Cancer with Prostate-Specific Antigen Testing: American Society of Clinical Oncology Provisional Clinical Opinion. J Clin Oncol 30(24):3020-5.
				Study type (Level)	SR (I)
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
	✓			• Were the databases searched reported?	III
		✓		• Was more than one database searched?	III
	✓			• Were search terms reported?	IV
		✓		• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
	✓			• Were inclusion/exclusion criteria reported?	II
		✓		• Was the inclusion criteria applied in an unbiased way?	III
	✓			• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
✓				• Was the quality of the studies reported?	III
✓				• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	
✓				• Were the characteristics of the individual studies reported?	III
✓				• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	
			✓	• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
			✓	• Was a test for heterogeneity applied?	III-IV
			✓	• If there was heterogeneity, was this discussed or the reasons explored?	III-IV
				Comments:	• The authors state that the literature search of the systematic review conducted by the AHRQ (Lin et al, 2011; Chou et al, 2011) was used as the basis of an update search to March 16, 2012. However, the databases searched and the search terms used are not specifically defined. It is also unknown if the literature search included hand searching. Inclusion/exclusion criteria were not reported and it is not known if the inclusion criteria were applied in an unbiased way. It is thus unclear if an adequate systematic search strategy was applied. • Lower levels of evidence were included to describe the adverse events associated with prostate biopsy. However, these results were reported individually and separately to the RCTs and SR evidence. • The quality of the studies was taken from the AHRQ systematic review using the quality appraisal methods of the USPSTF. The authors did not conduct a separate quality assessment. • No statistical analysis of data was conducted by the authors of the study.

Quality rating[b]:	Systematic review: Poor	
	Included Level II studies: 1) PLCO – Andriole et al (2009) and Andriole et al (2012) 2) ERSPC – Schroder et al (2009) 3) Stockholm – Kjellman et al (2009) 4) Quebec – Labrie et al (2004) 5) Norrkoping – Sandblom et al (2004) and Sandblom et al (2011) 6) Goteborg – Hugosson et al (2010)	

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; ERSPC, European Randomised Study of Screening for Prostate Cancer; RCT, randomised controlled trial; N, no; NA, not applicable; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; SR, systematic review; USPSTF, U.S. Preventive Services Task Force; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

			Study ID	**Djulbegovic et al (2010)**	
			Citation	Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, and Dahm P. (2010). Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. 2010. BMJ 341:c4543.	
			Study type (Level)	SR MA (I)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
✓				• Were the databases searched reported?	III
✓				• Was more than one database searched?	III
✓				• Were search terms reported?	IV
✓				• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
✓				• Were inclusion/exclusion criteria reported?	II
✓				• Was the inclusion criteria applied in an unbiased way?	III
✓				• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
✓				• Was the quality of the studies reported?	III
✓				• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	
✓				• Were the characteristics of the individual studies reported?	III
✓				• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	
✓				• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
✓				• Was a test for heterogeneity applied?	III-IV
✓				• If there was heterogeneity, was this discussed or the reasons explored?	III-IV

Comments:	• The quality of the included studies was assessed using the GRADE criteria. Whilst an overall quality rating of the individual RCTs was not reported, the quality rating of the individual components of the GRADE criteria was presented. The overall quality of evidence/GRADE result for each outcome that was meta-analysed was also shown. • Relative risks were used to summarise the effect of screening intervention for all outcomes. Mantel-Haenszel estimates were calculated based on the number of participants in a given study arm and pooled under a random effects model, with data expressed as relative risks and 95% confidence intervals. When no information on event rates was available, the inverse variance method was used.	
Quality rating[b]:	Systematic review: Good	
	Included Level II studies: 1) Quebec – Labrie et al (1988), Labrie et al (1999) 2) Norrkoping – Sandblom et al (2004), Varenhorst et al (1992) 3) ERSPC – Schroder et al (2009) 4) French ERSPC – Jegu et al (2009) 5) PLCO – Andriole et al (2009) 6) Goteborg – Hugosson et al (2010)	

Abbreviations: GRADE, Grading of Recommendations Assessment, Development and Evaluation; ERSPC, European Randomised Study of Screening for Prostate Cancer; MA, meta-analysis; N, no; NA, not applicable; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; SR, systematic review; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

Study ID				**Hamashima et al (2009)**	
Citation				Hamashima C, Nakayama T, Sagawa M, Saito H, Sobue T (2009) The Japanese guideline for prostate cancer screening. Jpn J Clin Oncol 39(6):339-51.	
Study type (Level)				SR MA (I)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
✓				• Were the databases searched reported?	III
✓				• Was more than one database searched?	III
	✓			• Were search terms reported?	IV
✓				• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
	✓			• Were inclusion/exclusion criteria reported?	II
		✓		• Was the inclusion criteria applied in an unbiased way?	III
	✓			• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
	✓			• Was the quality of the studies reported?	III
		✓		• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	
✓				• Were the characteristics of the individual studies reported?	III
✓				• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	
			✓	• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
	✓			• Was a test for heterogeneity applied?	III-IV
			✓	• If there was heterogeneity, was this discussed or the reasons explored?	III-IV

Comments:	• The general systematic search strategy was outlined. However, search terms were not reported, nor the inclusion/exclusion criteria. Hence, this automatically led to a poor rating as it is unclear if an adequate systematic search strategy was applied. • Lower levels of evidence (e.g. cohort and case-control studies) were included in the SR. However, the results for each study type were reported separately. • The authors did not formally assess the quality of the included studies but the quality of the evidence was narratively discussed in the text. A clear-pre-determined strategy to assess study quality was not reported. • No statistical analysis of data was conducted by the authors of the study. • A full version of this report is available online in Japanese at the following website: http://canscreen.ncc.go.jp/	
Quality rating[b]:	Systematic review: Poor	
	Included Level II studies: 1) Quebec – Labrie et al (2009) 2) Norrkoping – Sandblom et al (2004) 3) Swedish ERSPC – Aus et al (2007)	

Abbreviations: MA, meta-analysis; N, no; NA, not applicable; NR, not reported; SR, systematic review; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

Study ID				**Ilic et al (2013) [Cochrane review]**	
Citation				Ilic D, Neuberger MM, Djulbegovic M, and Dahm P. (2013). Screening for prostate cancer. Cochrane Database of Systematic Reviews. 2006 Issue 3. Art. No.: CD004720. DOI: 10.1002/14651858.CD004720.pub2.	
Study type (Level)				SR MA (I)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
✓				• Were the databases searched reported?	III
✓				• Was more than one database searched?	III
✓				• Were search terms reported?	IV
✓				• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
✓				• Were inclusion/exclusion criteria reported?	II
✓				• Was the inclusion criteria applied in an unbiased way?	III
✓				• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
✓				• Was the quality of the studies reported?	III
✓				• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	
✓				• Were the characteristics of the individual studies reported?	III
✓				• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	
✓				• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
✓				• Was a test for heterogeneity applied?	III-IV
✓				• If there was heterogeneity, was this discussed or the reasons explored?	III-IV
Comments:				• The quality of the studies was assessed by The Cochrane Collaboration's tool for	

	assessing risk of bias, as described in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011). • Additionally, the GRADE framework was applied to rate the quality of evidence for each outcome. • Data were analysed according to intention-to-screen analysis. This included re-analysing the results from the Quebec trial using intention-to-screen analysis according to the groups to which the participants were originally randomised (i.e. screening versus control).	
Quality rating[b]:	Systematic review: Good	
	Included Level II studies: 1) ERSPC – Schroder et al (2009) 2) Norrkoping – Sandblom et al (2004) 3) PLCO – Andriole et al (2009) 4) Quebec – Labrie et al (1988) 5) Stockholm – Kjellman et al (2009)	

Abbreviations: ERSPC, European Randomised Study of Screening for Prostate Cancer; GRADE, Grading of Recommendations Assessment, Development and Evaluation; MA, meta-analysis; N, no; NA, not applicable; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; SR, systematic review; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

Study ID				**Lin et al (2011) [AHRQ]**	
Citation				Lin K, Croswell JM, Koenig H, Lam C, and Maltz A. (2011). Prostate-specific antigen-based screening for prostate cancer: an evidence update for the U.S. preventive services task force. Evidence synthesis no. 90. AHRQ publication no. 12-05160-EF-1. Rockville, MD: Agency for Healthcare Research and Quality. Associated publication: Chou R, Croswell JM, Dana T, Bougatsos C, Blazina I, Fu R, Gleitsmann K, Koenig HC, Lam C, Maltz A, Rugge JB and Lin K. (2011). Screening for prostate cancer: a review of the evidence for the U.S. preventative services task force. Ann Intern Med 155:762-771.	
Study type (Level)				SR (I)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
✓				• Were the databases searched reported?	III
✓				• Was more than one database searched?	III
✓				• Were search terms reported?	IV
✓				• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
✓				• Were inclusion/exclusion criteria reported?	II
✓				• Was the inclusion criteria applied in an unbiased way?	III
✓				• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
✓				• Was the quality of the studies reported?	III
✓				• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	
✓				• Were the characteristics of the individual studies reported?	III
	✓			• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	

			✓	• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
			✓	• Was a test for heterogeneity applied?	III-IV
			✓	• If there was heterogeneity, was this discussed or the reasons explored?	III-IV
Comments:				• The authors used predefined U.S. Preventative Services Task force criteria to rate the quality of the included studies. • Only PubMed and Cochrane Database were searched. • No meta-analysis was completed by the authors of the SR. Rather, the results of the individual included studies were discussed and a descriptive overall conclusion was drawn by the authors. • Sources of heterogeneity were not explored. • No statistical analysis of data was conducted by the authors of the study.	
Quality rating[b]:				Systematic review: Good	
				Included Level II studies: 1) Schroder et al, 2009 (ERSPC) 2) Andriole et al, 2009 (PLCO) 3) Hugosson et al, 2010 (Goteborg) 4) Kjellmen et al, 2009 (Stockholm) 5) Sandblom et al, 2011 (Norrkoping) Two meta-analyses were also included: 1) Djulbegovic et al, 2010 2) Ilic et al, 2011	

Abbreviation: AHRQ, Agency for Healthcare Research and Quality; N, no; NA, not applicable; NR, not reported; SR, systematic review; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

Study ID				**Lumen et al (2012)**	
Citation				Lumen N, Fonteyne V, de Meerleert G, Ost P, Villeirs G, Mottrie A, de Visschere P, de Troyer B, and Oosterlinck, W. (2012). Population screening for prostate cancer: an overview of available studies and meta-analysis. Int J Urol 19:100-108.	
Study type (Level)				SR MA (I) The authors included the Rotterdam-Ireland trial which it notes was not a prospective RCT. Rather, it was a comparison between a screened population (part of the Rotterdam section of the ERSPC trial) and a population where screening is not routinely carried out (Ireland). Consequently, the evidence reviewer acknowledges that Lumen (2012) does not fit precisely into NHMRC's classification of a Level I study.	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
✓				• Were the databases searched reported?	III
✓				• Was more than one database searched?	III
✓				• Were search terms reported?	IV
		✓		• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
✓				• Were inclusion/exclusion criteria reported?	II
✓				• Was the inclusion criteria applied in an unbiased way?	III
	✓			• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
	✓			• Was the quality of the studies reported?	III
			✓	• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	

✓				• Were the characteristics of the individual studies reported?	III
✓				• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	
✓				• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
✓				• Was a test for heterogeneity applied?	III-IV
✓				• If there was heterogeneity, was this discussed or the reasons explored?	III-IV
Comments:				• All of the included studies were RCTs with the exception of the Rotterdam-Ireland trial which was described as follows: "Not a prospective randomized clinical trial, but a comparison between a screened population (part of the Rotterdam section of the ERSPC) and a population where screening is not routinely carried out (Ireland)." Hence, the evidence reviewer acknowledges that Lumen (2012) does not fit precisely into NHMRC's classification of a Level I study. • Whilst the quality of the individual studies was not reported, the authors reference the extensive quality assessment of the individual studies with evaluation of the potential sources of bias that were identified by two other recent SRs (Djulbegobic et al, 2010; Ilic et al, 2011). Consequently, the authors did not conduct a separate quality assessment in this SR. • Data were analysed according to intention-to-screen analysis. This included re-analysing the results from the Quebec trial using intention-to-screen analysis according to the groups to which the participants were originally randomised (i.e. screening versus control).	
Quality rating[b]:				Systematic review: Good	
				Included Level II studies: 1) Norrkoping – Sandblom et al (2004); Sandblom et al (2011) 2) Quebec – Labrie et al (2004); Labrie et al (1999) 3) ERSPC – Schroder et al (2002); Roobol et al (2009) 4) PLCO – Andriole et al (2009) 5) Goteborg – Hugosson et al (2010); Aus et al (2007) 6) Rotterdam-Ireland – Van Leeuwen et al (2010) 7) French ERSPC – Jegu et al (2009) 8) Stockholm – Kjellman et al (2009)	

Abbreviations: MA, meta-analysis; N, no; NA, not applicable; NHMRC, National Health and Medical Research Council; NR, not reported; RCT, randomised controlled trial; SR, systematic review; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

Study ID				**NZGG (2009)**	
Citation				New Zealand Guidelines Group. (2009). Cancer control strategy guidance completion: update of evidence for prostate-specific antigen (PSA) testing in asymptomatic men. Wellington: Ministry of Health.	
Study type (Level)				SR (I)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was an adequate search strategy used?	
✓				• Was a systematic search strategy reported?	I
✓				• Were the databases searched reported?	III
✓				• Was more than one database searched?	III
✓				• Were search terms reported?	IV
✓				• Did the literature search include hand searching?	IV
				B. Were the inclusion criteria appropriate and applied in an unbiased way?	
✓				• Were inclusion/exclusion criteria reported?	II

✓				• Was the inclusion criteria applied in an unbiased way?	III
	✓			• Was only level II evidence included?	I-IV
				C. Was a quality assessment of included studies undertaken?	
✓				• Was the quality of the studies reported?	III
✓				• Was a clear, pre-determined strategy used to assess study quality?	IV
				D. Were the characteristics and results of the individual studies appropriately summarised?	
✓				• Were the characteristics of the individual studies reported?	III
✓				• Were baseline demographic and clinical characteristics reported for patients in the individual studies?	IV
✓				• Were the results of the individual studies reported?	III
				E. Were the methods for pooling the data appropriate?	
			✓	• If appropriate, was a meta-analysis conducted?	III-IV
				F. Were the sources of heterogeneity explored?	
			✓	• Was a test for heterogeneity applied?	III-IV
			✓	• If there was heterogeneity, was this discussed or the reasons explored?	III-IV
Comments:				• Used adapted checklists from the GATE framework to evaluate the quality of Level I-IV studies. • Case series are summarised in evidence tables but were not formally appraised using a critical appraisal checklist. • No meta-analysis was completed by the authors of the SR. Rather, the results of the individual included studies were discussed and a descriptive overall conclusion was drawn by the authors. • Whilst all levels of evidence were included, Level I and II evidence were reported separately to lower levels of evidence. • No statistical analysis of data was conducted by the authors of the study.	
Quality rating[b]:				Systematic review: Good	
				Included Level II studies: Refer to data extraction form for Level II studies included for each outcome.	

Abbreviations: GATE, Graphic Appraisal Tool for Epidemiology; N, no; NA, not applicable; NR, not reported; SR, systematic review; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings applied to the systematic reviews are good, fair or poor. The quality of the individual Level II studies included in the systematic reviews are recorded as reported in the systematic reviews.

1.7.2 Randomised controlled trials

			Study ID	**PLCO**	
			Citation	Andriole GL, Crawford DE, Grubb III RL, Buys SS, Chia D, Church TR, Fouad MN, Isaacs C et al. (2012). Prostate cancer screening in the randomised prostate, lung, colorectal, and ovarian cancer screening trial: mortality results after 13 years of follow-up. J Natl Cancer Inst 104:125-132.	
			Study type (Level)	RCT (II)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was assignment of subjects to treatment group randomised?	
✓				• Was the use of randomisation reported?	I
✓				• Was the method of randomisation reported?	III
✓				• Was the method of randomisation appropriate?	I-III
				A. Was allocation to treatment groups concealed from those responsible for recruiting subjects?	
✓				• Was a method of allocation concealment reported?	III
✓				• Was the method of allocation concealment adequate?	III
				B. Was the study double-blinded?	
	✓			• Were subjects and investigators blinded to treatment arm?	II-IV
				C. Were patient characteristics and demographics similar between treatment arms at baseline?	
✓				• Were baseline patient characteristics and demographics reported?	III
✓				• Were the characteristics similar between treatment arms?	III-IV
				D. Were all randomised participants included in the analysis?	
✓				• Was loss to follow-up reported?	II
✓				• Was loss to follow-up appropriately accounted for in the analysis?	III-IV
				E. Was outcome assessment likely to be subject to bias?	
✓				• Were all relevant outcomes measured in a standard, valid, and reliable way?	III-IV
✓				• Was outcome assessment blinded to treatment allocation?	III
			✓	• If outcome assessment was not blinded, were outcomes objective and unlikely to be influenced by blinding of assessment?	III
				F. Were the statistical methods appropriate?	
✓				• Were the methods used for comparing results between treatment arms appropriate?	III
✓				• If the study was carried out at more than one site, are the results comparable for all sites?	IV
				G. If appropriate, were any subgroup analyses carried out?	
✓				• Were subgroup analyses reported?	III-IV
✓				• Were subgroup analyses appropriate?	III-IV
			Comments:	• Allocation concealment was achieved through the use of a central system (Prorok et al, 2000) • It is not possible to blind participants and clinicians to the screening intervention. Data on diagnosed cancers and mortality were obtained by patient reported questionnaire and followed up by telephone (unblinded). Possible cancer-specific deaths were reviewed by blinded reviewers • A detailed list of baseline patient characteristics and demographics were reported, including age, race or ethnic group, family history of prostate cancer, and previous PSA or DRE tests • Data were analysed by intention-to-screen analysis • Contamination rate: increased from 40% in the 1st year to 52% in the 6th year of PSA testing	

	• Compliance rate for screening: 85% for PSA and 86% for DRE	
Quality rating[b]:	Good	

Abbreviations: N, no; NA, not applicable; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; RCT, randomised controlled trial; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra. The quality of the individual Level II studies included in the systematic reviews will not be assessed as part of this evidence review. However, the quality assessment forms for the level I evidence will capture the quality of the individual Level II studies, as reported.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings are good, fair or poor. The quality rating should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

			Study ID	**ERSPC**	
			Citation	Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M et al. (2012). Prostate cancer mortality at 11 years of follow-up. N Eng J Med 366:981-990.	
			Study type (Level)	RCT (II)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was assignment of subjects to treatment group randomised?	
✓				• Was the use of randomisation reported?	I
✓				• Was the method of randomisation reported?	III
✓				• Was the method of randomisation appropriate?	I-III
				A. Was allocation to treatment groups concealed from those responsible for recruiting subjects?	
		✓		• Was a method of allocation concealment reported?	III
			✓	• Was the method of allocation concealment adequate?	III
				B. Was the study double-blinded?	
	✓			• Were subjects and investigators blinded to treatment arm?	II-IV
				C. Were patient characteristics and demographics similar between treatment arms at baseline?	
	✓			• Were baseline patient characteristics and demographics reported?	III
		✓		• Were the characteristics similar between treatment arms?	III-IV
				D. Were all randomised participants included in the analysis?	
✓				• Was loss to follow-up reported?	II
✓				• Was loss to follow-up appropriately accounted for in the analysis?	III-IV
				E. Was outcome assessment likely to be subject to bias?	
✓				• Were all relevant outcomes measured in a standard, valid, and reliable way?	III-IV
✓				• Was outcome assessment blinded to treatment allocation?	III
			✓	• If outcome assessment was not blinded, were outcomes objective and unlikely to be influenced by blinding of assessment?	III
				F. Were the statistical methods appropriate?	
✓				• Were the methods used for comparing results between treatment arms appropriate?	III
✓				• If the study was carried out at more than one site, are the results comparable for all sites?	IV
				G. If appropriate, were any subgroup analyses carried out?	
✓				• Were subgroup analyses reported?	III-IV
✓				• Were subgroup analyses appropriate?	III-IV
			Comments:	• Method of allocation concealment not reported.	

	• It is not possible to blind participants and clinicians to the screening intervention. However, causes of death were evaluated in a blinded manner. • The only baseline patient characteristics and demographics reported were the mean and median age of participants at randomisation, with "little variation among the seven countries". Mean (and median) age was similar between study arms. • Withdrawals of two participating study sites were not included due to short duration of follow-up and discontinuation in the whole ERSPC study. • Data were analysed by intention-to-screen analysis. • Contamination rate: estimated to be 30.7% (Roobol et al, 2009) • Compliance rate for screening: 82.2%	
Quality rating[b]:	Fair	

Abbreviations: ERSPC, European Randomised Study of Screening for Prostate Cancer; N, no; NA, not applicable; NR, not reported; RCT, randomised controlled trial; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra. The quality of the individual Level II studies included in the systematic reviews will not be assessed as part of this evidence review. However, the quality assessment forms for the level I evidence will capture the quality of the individual Level II studies, as reported.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings are good, fair or poor. The quality rating should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

			Study ID	**Goteborg**	
			Citation	Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, Pihl CG, Stranne J et al. (2010). Mortality results from the Goteborg randomised population-based prostate cancer screening trial. Lancet Oncol 11:725-32	
			Study type/level	RCT (II)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was assignment of subjects to treatment group randomised?	
✓				• Was the use of randomisation reported?	I
✓				• Was the method of randomisation reported?	III
✓				• Was the method of randomisation appropriate?	I-III
				A. Was allocation to treatment groups concealed from those responsible for recruiting subjects?	
		✓		• Was a method of allocation concealment reported?	III
			✓	• Was the method of allocation concealment adequate?	III
				B. Was the study double-blinded?	
	✓			• Were subjects and investigators blinded to treatment arm?	II-IV
				C. Were patient characteristics and demographics similar between treatment arms at baseline?	
	✓			• Were baseline patient characteristics and demographics reported?	III
		✓		• Were the characteristics similar between treatment arms?	III-IV
				D. Were all randomised participants included in the analysis?	
✓				• Was loss to follow-up reported?	II
✓				• Was loss to follow-up appropriately accounted for in the analysis?	III-IV
				E. Was outcome assessment likely to be subject to bias?	
✓				• Were all relevant outcomes measured in a standard, valid, and reliable way?	III-IV
✓				• Was outcome assessment blinded to treatment allocation?	III
			✓	• If outcome assessment was not blinded, were outcomes objective and unlikely to be influenced by blinding of assessment?	III

				F. Were the statistical methods appropriate?	
✓				• Were the methods used for comparing results between treatment arms appropriate?	III
			✓	• If the study was carried out at more than one site, are the results comparable for all sites?	IV
				G. If appropriate, were any subgroup analyses carried out?	
✓				• Were subgroup analyses reported?	III-IV
✓				• Were subgroup analyses appropriate?	III-IV
Comments:				• According to the authors, the randomisation procedure was done at the Department of Statistics at the University of Göteborg. 10-digit personal identifiers were the only available personal data for those doing the computer randomisation. Although not stated, the personal identifiers may have been Swedish national ID numbers. • It is not possible to blind participants and clinicians to the screening intervention. • No baseline sociodemographic comparison of the two groups. Only the age of the participants was reported. • Data were analysed according to the intention-to-screen principle. • Contamination rate: not reported. Only described as "low". • Compliance rate for screening: 76%	
Quality rating[b]:				Fair	

Abbreviations: N, no; NA, not applicable; NR, not reported; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra. The quality of the individual Level II studies included in the systematic reviews will not be assessed as part of this evidence review. However, the quality assessment forms for the level I evidence will capture the quality of the individual Level II studies, as reported.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings are good, fair or poor. The quality rating should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

Study ID				**Norrkoping**	
Citation				Sandblom G, Varenhorst E, Rosell J, Lofman O and Carlsson P. (2011). Randomised prostate cancer screening trial: 20 year follow-up. BMJ 342: d1539.	
Study type (Level)				Psuedo-RCT (III-1)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was assignment of subjects to treatment group randomised?	
✓				• Was the use of randomisation reported?	I
✓				• Was the method of randomisation reported?	III
	✓			Was the method of randomisation appropriate?	I-III
				A. Was allocation to treatment groups concealed from those responsible for recruiting subjects?	
	✓			• Was a method of allocation concealment reported?	III
			✓	• Was the method of allocation concealment adequate?	III
				B. Was the study double-blinded?	
	✓			• Were subjects and investigators blinded to treatment arm?	II-IV
				C. Were patient characteristics and demographics similar between treatment arms at baseline?	
	✓			• Were baseline patient characteristics and demographics reported?	III
		✓		• Were the characteristics similar between treatment arms?	III-IV
				D. Were all randomised participants included in the analysis?	
✓				• Was loss to follow-up reported?	II
✓				• Was loss to follow-up appropriately accounted for in the analysis?	III-IV

				E. Was outcome assessment likely to be subject to bias?	
✓				• Were all relevant outcomes measured in a standard, valid, and reliable way?	III-IV
	✓			• Was outcome assessment blinded to treatment allocation?	III
✓				• If outcome assessment was not blinded, were outcomes objective and unlikely to be influenced by blinding of assessment?	III
				F. Were the statistical methods appropriate?	
✓				• Were the methods used for comparing results between treatment arms appropriate?	III
			✓	• If the study was carried out at more than one site, are the results comparable for all sites?	IV
				G. If appropriate, were any subgroup analyses carried out?	
	✓			• Were subgroup analyses reported?	III-IV
			✓	• Were subgroup analyses appropriate?	III-IV
Comments:				• The method of randomisation by allocating every 6th man to the screening group from a list of date of births is a predictable group assignment and hence this study is classified as Level III-1 evidence (pseudo-RCT). • There was no description of allocation concealment. • It is not possible to blind participants and clinicians to the screening intervention. There was no specific mention of blinding of outcome assessors; however outcomes and outcome measurements are unlikely to be influenced by lack of blinding. • No baseline sociodemographic comparison of the two groups. Only the age of the participants was reported. • Withdrawals were cited but it is unclear how the data for those men who migrated was available. • Data were analysed by intention-to-screen analysis. • Contamination rate: not specified but the authors report on a "low rate of contamination". • Compliance rate for screening: 70-78% depending on year	
Quality rating[b]:				Poor	

Abbreviations: N, no; NA, not applicable; NR, not reported; RCT, randomised controlled trial; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra. The quality of the individual Level II studies included in the systematic reviews will not be assessed as part of this evidence review. However, the quality assessment forms for the level I evidence will capture the quality of the individual Level II studies, as reported.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings are good, fair or poor. The quality rating should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

Study ID				**Stockholm**	
Citation				Kjellman A, Akre O, Norming U, Tornblom M, and Gustafsson O. (2009). 15-year followup of a population based prostate cancer screening study. J Urol 181:1615-1621.	
Study type (Level)				RCT (II)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was assignment of subjects to treatment group randomised?	
✓				• Was the use of randomisation reported?	I
	✓			• Was the method of randomisation reported?	III
			✓	• Was the method of randomisation appropriate?	I-III
				A. Was allocation to treatment groups concealed from those responsible for recruiting subjects?	
	✓			• Was a method of allocation concealment reported?	III
			✓	• Was the method of allocation concealment adequate?	III

				B. Was the study double-blinded?	
	✓			• Were subjects and investigators blinded to treatment arm?	II-IV
				C. Were patient characteristics and demographics similar between treatment arms at baseline?	
	✓			• Were baseline patient characteristics and demographics reported?	III
		✓		• Were the characteristics similar between treatment arms?	III-IV
				D. Were all randomised participants included in the analysis?	
✓				• Was loss to follow-up reported?	II
✓				• Was loss to follow-up appropriately accounted for in the analysis?	III-IV
				E. Was outcome assessment likely to be subject to bias?	
✓				• Were all relevant outcomes measured in a standard, valid, and reliable way?	III-IV
	✓			• Was outcome assessment blinded to treatment allocation?	III
✓				• If outcome assessment was not blinded, were outcomes objective and unlikely to be influenced by blinding of assessment?	III
				F. Were the statistical methods appropriate?	
✓				• Were the methods used for comparing results between treatment arms appropriate?	III
			✓	• If the study was carried out at more than one site, are the results comparable for all sites?	IV
				G. If appropriate, were any subgroup analyses carried out?	
	✓			• Were subgroup analyses reported?	III-IV
			✓	• Were subgroup analyses appropriate?	III-IV
Comments:				• Methods of randomisation and allocation concealment were not described. • It is not possible to blind participants and clinicians to the screening intervention. There was no specific mention of blinding of outcome assessors; however outcomes and outcome measurement are unlikely to be influenced by lack of blinding. • No baseline sociodemographic comparison of the two groups. Only the age of the participants was reported. • Whilst there was no missing outcome data for mortality or number diagnosed, the report has internal discrepancies about the total number of participants because the file containing the registration numbers of the original cohort could not be retrieved. • Data were analysed by intention-to-screen analysis. • Contamination rate: not reported • Compliance rate for screening: 74%	
Quality rating[b]:				Poor	

Abbreviations: N, no; NA, not applicable; NR, not reported; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra. The quality of the individual Level II studies included in the systematic reviews will not be assessed as part of this evidence review. However, the quality assessment forms for the level I evidence will capture the quality of the individual Level II studies, as reported.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings are good, fair or poor. The quality rating should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

Study ID				**Quebec**	
Citation				Labrie F, Candas B, Cusan L, Gomez JL, Belanger A, Brousseau g, Chebrette E and Levesque J. (2004). Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. Prostate 59(3): 311-318.	
Study type (Level)				RCT (II)	
Y	**N**	**NR**	**NA**	**Quality criteria**	**Error category[a]**
				A. Was assignment of subjects to treatment group randomised?	
✓				• Was the use of randomisation reported?	I
	✓			• Was the method of randomisation reported?	III
			✓	• Was the method of randomisation appropriate?	I-III
				A. Was allocation to treatment groups concealed from those responsible for recruiting subjects?	
	✓			• Was a method of allocation concealment reported?	III
			✓	• Was the method of allocation concealment adequate?	III
				B. Was the study double-blinded?	
	✓			• Were subjects and investigators blinded to treatment arm?	II-IV
				C. Were patient characteristics and demographics similar between treatment arms at baseline?	
	✓			• Were baseline patient characteristics and demographics reported?	III
		✓		• Were the characteristics similar between treatment arms?	III-IV
				D. Were all randomised participants included in the analysis?	
✓				• Was loss to follow-up reported?	II
✓				• Was loss to follow-up appropriately accounted for in the analysis?	III-IV
				E. Was outcome assessment likely to be subject to bias?	
✓				• Were all relevant outcomes measured in a standard, valid, and reliable way?	III-IV
	✓			• Was outcome assessment blinded to treatment allocation?	III
✓				• If outcome assessment was not blinded, were outcomes objective and unlikely to be influenced by blinding of assessment?	III
				F. Were the statistical methods appropriate?	
	✓			• Were the methods used for comparing results between treatment arms appropriate?	III
			✓	• If the study was carried out at more than one site, are the results comparable for all sites?	IV
				G. If appropriate, were any subgroup analyses carried out?	
	✓			• Were subgroup analyses reported?	III-IV
			✓	• Were subgroup analyses appropriate?	III-IV
Comments:				• Methods of randomisation and allocation concealment were not described. • It is not possible to blind participants and clinicians to the screening intervention. Blinding of outcome assessment was not clearly described. • No baseline sociodemographic comparison of the two groups. Only the age of the participants was reported. • Withdrawals from both the screening and control groups were cited. • Data were not analysed according to the intention-to-screen analysis. Rather, due to the high crossover rate, the authors of the study decided to analyse the data according to whether the participants actually received screening or not. • Contamination rate: 7.3% • Compliance rate for screening: 23.6%	
Quality rating[b]:				Poor	

Abbreviations: N, no; NA, not applicable; NR, not reported; RCT, randomised controlled trial; Y, yes.

Note: Quality criteria adapted from NHMRC (2000) How to use the evidence: assessment and application of scientific evidence. NHMRC, Canberra. The quality of the individual Level II studies included in the systematic reviews will not be assessed as part of this evidence review. However, the quality assessment forms for the level I evidence will capture the quality of the individual Level II studies, as reported.

[a] Error categories as follows: (I) leads to exclusion of the study; (II) automatically leads to a poor rating; (III) leads to a one grade reduction in quality rating (eg, good to fair, or fair to poor); and (IV) errors that are may or may not be sufficient to lead to a decrease in rating.

[b] Quality ratings are good, fair or poor. The quality rating should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

1.8 Data extraction forms

1.8.1 Systematic reviews

Study ID	**Ilic et al (2013) [Cochrane review]**
Citation/Primary publication	Ilic D, Neuberger MM, Djulbegovic M, and Dahm P. (2013). Screening for prostate cancer. Cochrane Database of Systematic Reviews. 2013 Issue 1. Art. No.: CD004720. DOI: 10.1002/14651858.CD004720.pub3.
Study type (Level)	SR MA (I)
Affiliation/Source of funds	• Department of Urology, College of Medicine, University of Florida, USA • Malcolm Randall Veterans Affairs Medical Centre, Gainesville, Florida, USA • Dennis W. Jahnigen Career Development Scholars Award by the American Geriatrics Society, USA
Intervention	
Description of test +/- DRE +/- TRUS	Studies that used any of the following screening measures, individually or in combination were included: • DRE • PSA test (including total, velocity, density and percentage free and complex) • TRUS guided biopsy
PSA test cut-off	Not specifically defined – dependent on the PSA test cut-off used in each individual study
Comparator	
Description of comparator	No screening for prostate cancer
Eligibility criteria	
Inclusion criteria	All men enrolled in studies of prostate cancer screening with no exclusions based on ethnicity, age or presence of LUTS
Definition of asymptomatic	Not specifically defined
Exclusion criteria	Men with a previous diagnosis and treatment of prostate cancer
Literature search	
Search strategy Search period? Publication types? English only?	• Search period - Original 2006 version of this review: PROSTATE register was searched in November 2004 and the remaining databases were searched for studies published between 1966 and January 2006 - Current 2012 version of this review: an updated search of the electronic databases was performed with the existing search strategy in June 2012 • Publication types: all RCTs and quasi-RCTs of screening versus no screening for prostate cancer • No language restrictions were placed on studies considered for inclusion • Published or unpublished sources were considered
Exclusion criteria	See above

Databases searched	• PROSTATE register, CENTRAL (Cochrane Central Register of Controlled Trials), MEDLINE, EMBASE, CANCERLIT and the NHS EED • Hand searching for reviews and technical reports with regard to prostate cancer screening in specialist journals and grey literature • The following journals were hand searched until March 2005 - BJU International (2000-2005) - European Urology (2002-2005) - The Prostate (1998-2005) - Journal of Urology (1996-2005) - Urology (2002-2005) - Cancer (1998-2005) • Abstracts from the following meetings were manually searched from 2005-2012 - American Urological Association - European Association of Urology - American Society of Clinical Oncology				
Outcomes and measures					
Primary outcome	The efficacy of screening men for prostate cancer in reducing prostate cancer-specific and all-cause mortality				
Secondary outcomes	• Incident prostate cancers by stage and grade at diagnosis • Metastatic disease at follow-up • Quality of life • Harms of screening (including both adverse outcomes from false-positive and/or false negative results and their impact upon resulting treatment procedures • Costs associated with screening programs				
Other outcomes	None				
Key results					
Outcome	**No. trials (no. patients)**	**Screening n/N (%)**	**Control n/N (%)**	**Relative risk estimate (95% CI)**	**P-value** • *Favours screening/control or no difference* • *Substantial /moderate/mild heterogeneity[a] P=X (I2=X)*
Prostate cancer-specific mortality					
Sensitivity analysis: risk of bias					
Low risk of bias studies ERSPC PLCO	2 (258,684)	462/121,156 (0.38)	607/137,528 (0.44)	0.96 (0.70-1.30)	• No difference P=0.77 • Substantial heterogeneity P=0.05 (I2=74%)
Total	2 (258,684)	462/121,156 (0.38)	607/137,528 (0.44)	0.96 (0.70-1.30)	• No difference P=0.77 • Substantial heterogeneity P=0.05 (I2=74%)
Sensitivity analysis: risk of bias, including ERSPC's core age group consisting of men aged 55-69 years					
Low risk of bias studies ERSPC PLCO	2 (238,928)	397/111,231 (0.36)	547/127,697 (0.43)	0.94 (0.65-1.35)	• No difference P=0.72 • Substantial heterogeneity P=0.02 (I2=80%)
Total	2 (238,928)	397/111,231 (0.36)	547/127,697	0.94 (0.65-1.35)	• No difference P=0.72 • Substantial heterogeneity P=0.02 (I2=80%)
Subgroup analysis: age					
Men aged ≥45 years Quebec	1 (46,486)	153/31,133 (0.49)	75/15,353 (0.49)	1.01 (0.76-1.33)	• No difference P=0.97 • Heterogeneity: N/A

Men aged ≥50 years ERSPC Norrkoping	2 (191,025)	394/84,310 (0.47)	652/106,715 (0.61)	0.93 (0.69-1.27)	• No difference P=0.66 • Substantial heterogeneity P=0.12 (I2=59%)
Men aged ≥55 years PLCO Stockholm	2 (103,831)	151/40,714 (0.37)	591/63,117 (0.94)	1.12 (0.92-1.37)	• No difference P=0.26 • No significant heterogeneity P=0.79 (I2=0%)
Total	5 (341,342)	698/156,157 (0.45)	1318/185,185 (0.71)	1.00 (0.86-1.17)	• No difference P=0.99 • Moderate heterogeneity P=0.12 (I2=46%)
Subgroup analysis: age, including ERSPC's core age group consisting of men aged between 55-69 years					
Men aged ≥45 years Quebec	1 (46,486)	153/31,133 (0.49)	75/15,353 (0.49)	1.01 (0.76-1.33)	• No difference P=0.97 • Heterogeneity: N/A
Men aged ≥50 years Norrkoping	1 (9026)	30/1494 (2.01)	130/7532 (1.73)	1.16 (0.79-1.72)	• No difference P=0.45 • Heterogeneity: N/A
Men aged ≥55 years ERSPC PLCO Stockholm	3 (266,074)	450/113,605 (0.40)	1053/152,469 (0.69)	0.98 (0.75-1.27)	• No difference P=0.86 • Substantial heterogeneity P=0.02 (I2=73%)
Total	5 (321,586)	633/146,232 (0.43)	1258/175,354 (0.72)	1.00 (0.83-1.19)	• No difference P=0.98 • Substantial heterogeneity P=0.05 (I2=58%)

Authors' comments

- The quality of evidence was rated as MODERATE for this outcome according to the GRADE approach
- The quality of the individual RCTs was assessed as:
  - Low risk of bias: ERSPC and PLCO
  - High risk of bias: Norrkoping, Quebec and Stockholm
- Participant characteristics including race/ethnicity, family history of prostate cancer, enlarged prostate (or BPH), previous prostate biopsy, PSA or DRE were only reported in the PLCO study
- The ERSPC study demonstrated a marginally significant benefit for screening in reducing prostate cancer-specific mortality among a 'core' subgroup of men aged 55-69 years at baseline (RR: 0.79; 95% CI 0.69-0.92) through a median follow-up duration of 11 years
- The PLCO study demonstrated no significant benefit for screening through 10 years of follow-up (RR: 1.15; 95% CI 0.86-1.54)
- Sensitivity analysis demonstrated no significant difference in results with the inclusion/exclusion of the Stockholm study

All-cause mortality

Outcome	No. trials (no. patients)	Screening n/N (%)	Control n/N (%)	Relative risk estimate (95% CI)	P-value • *Favours screening/control or no difference* • *Substantial /moderate/mild heterogeneity[a] P=X (I2=X)*
Subgroup analysis: age					
Men aged ≥50 years ERSPC Norrkoping	2 (191,025)	16,806/84,310 (19.93)	20,278/106,715 (19.00)	1.14 (0.84-1.56)	• No difference P=0.40 • Substantial heterogeneity P=0.02 (I2=83%)
Men aged ≥55 years PLCO Stockholm	2 (103,831)	6027/40,714 (14.80)	15,512/63,117 (24.58)	0.98 (0.95-1.01)	• No difference P=0.19 • No significant heterogeneity P=0.44 (I2=0%)

Total	4 (294,856)	22,833/125,024 (18.26)	35,790/169,832 (21.07)	1.00 (0.96-1.03)	• No difference P=0.84 • Substantial heterogeneity P=0.05 (I2=62%)
Subgroup analysis: age, including ERSPC's core age group consisting of men aged between 55-69 years					
Men aged ≥50 years Norrkoping	1 (9026)	69/1494 (4.62)	252/7532 (3.35)	1.38 (1.06-1.79)	• Favours control P=0.02 • Heterogeneity: N/A
Men aged ≥55 years PLCO ERSPC Stockholm	3 (266,074)	19,944/113,605 (17.56)	32,768/152,469 (21.49)	0.99 (0.97-1.00)	• No difference P=0.10 • No significant heterogeneity P=0.67 (I2=0%)
Total	4 (275,100)	20,013/115,099 (17.39)	33,020/160,001 (20.64)	0.99 (0.96-1.03)	• No difference P=0.59 • Substantial heterogeneity P=0.07 (I2=58%)
Sensitivity analysis: risk of bias					
Low risk of bias studies PLCO ERSPC	2 (258,684)	21,778/121,156 (17.98)	25,210/137,528 (18.33)	0.99 (0.96-1.02)	• No difference P=0.51 • Moderate heterogeneity P=0.16 (I2=49%)
Total	2 (258,684)	21,778/121,156 (17.98)	25,210/137,528 (18.33)	0.99 (0.96-1.02)	• No difference P=0.51 • Moderate heterogeneity P=0.16 (I2=49%)
Sensitivity analysis: risk of bias, including ERSPC's core age group consisting of men aged 55-69 years					
Low risk of bias studies PLCO ERSPC	2 (238,928)	18,958/111,231 (17.04)	22,440/127,697 (17.57)	0.98 (0.97-1.00)	• No difference P=0.09 • No significant heterogeneity P=0.44 (I2=0%)
Total	2 (238,928)	18,958/111,231 (17.04)	22,440/127,697 (17.57)	0.98 (0.97-1.00)	• No difference P=0.09 • No significant heterogeneity P=0.44 (I2=0%)

Authors' comments

NB. All-cause data provided in the PLCO trial report does not include deaths from prostate, lung or colorectal cancers. This unpublished data was obtained by the Cochrane review through author contact

- The quality of evidence was rated as MODERATE for this outcome according to the GRADE approach
- The quality of the individual RCTs was assessed as:
  - The ERSPC and PLCO studies were assessed as a low risk of bias
  - The Stockholm and Norrkoping studies were graded as a high risk of bias
- Sensitivity analysis demonstrated that there was no significant difference in results with the inclusion/exclusion of the Stockholm and Norrkoping study

Diagnosis of prostate cancer (as determined by study)

Outcome	No. trials (no. patients)	Screening n/N (%)	Control n/N (%)	Relative risk estimate (95% CI)	P-value
					• *Favours screening/control or no difference* • *Substantial /moderate/mild heterogeneity[a] P=X (I2=X)*
Subgroup analysis: age					

Men aged ≥50 years ERSPC Norrkoping	2 (191,025)	8023/84,310 (9.52)	6276/106,715 (5.88)	1.59 (1.54-1.64)	• Favours control (fewer events) P<0.00001 • No significant heterogeneity P=0.51 (I2=0%)
Men aged ≥55 years PLCO Stockholm	2 (103,831)	3906/40,714 (9.59)	5260/63,117 (8.33)	1.12 (1.08-1.17)	• Favours control (fewer events) P<0.00001 • No significant heterogeneity P=0.77 (I2=0%)
Total	4 (294,856)	11,929/125,024 (9.54)	11,536/169,832 (6.79)	1.30 (1.02-1.65)	• Favours control (fewer events) P=0.03 • Substantial heterogeneity P<0.00001 (I2=98%)
	NB. Incorporating data from the French site of the ERSPC study resulted in no change in findings (RR 1.26; 95% CI 1.06-1.51)				
Subgroup analysis: age, including ERSPC's 'core' age group consisting of men aged between 55-69 years					
Men aged ≥50 years Norrkoping	1 (9026)	85/1494 (5.69)	292/7532 (3.88)	1.47 (1.16-1.86)	• Favours control (fewer events) P=0.001 • Heterogeneity: N/A
Men aged ≥55 years ERSPC PLCO Stockholm	3 (266,074)	10,869/113,605 (9.57)	10,656/152,469 (6.99)	1.26 (0.96-1.64)	• No difference P=0.10 • Substantial heterogeneity P<0.00001 (I2=99%)
Total	4 (275,100)	10,954/115,099 (9.52)	10,948/160,001 (6.84)	1.30 (1.03-1.64)	• Favours control (fewer events) P=0.03 • Substantial heterogeneity P<0.00001 (I2=98%)
	A fixed effects model for the meta-analysis demonstrated no significant difference in results (RR 1.40; 95% CI 1.37-1.44)				

Authors' comments:

Bias

- The quality of evidence was rated as LOW for this outcome according the GRADE approach
- The quality of the individual RCTs was assessed as:
  - The ERSPC and PLCO studies were assessed as a low risk of bias
  - The Norrkoping and Stockholm studies were graded as a high risk of bias
- Sensitivity analysis demonstrated no meaningful differences in results with the exclusion of the Norrkoping and Stockholm studies

Heterogeneity

- Statistical heterogeneity was HIGH for this outcome
- Clinical heterogeneity was apparent with the Stockholm study as the screening procedures adopted in that study differed considerably from the other included studies. Sensitivity analysis demonstrated that there was no significant difference in results with the inclusion/exclusion of the Stockholm study
- Significant heterogeneity was associated with the meta-analysis for prostate cancer diagnosis. Performing a meta-analysis only according to age group significantly reduced the heterogeneity

Other comments

- In the ERSPC study, a total of 16.6% of screening tests were assessed as positive in the core age group, with 85.9% of men with positive tests undergoing a biopsy
- In the PLCO study, a total of 7.5% of men tested positive for a DRE and 7.9% for a PSA test, with 74% undertaking further diagnostic evaluation and 31.5% of men undergoing a biopsy within one year of screening

Prostate tumour stage					
Outcome	**No. trials (no. patients)**	**Screening n/N (%)**	**Control n/N (%)**	**Relative risk estimate (95% CI)**	**P-value** • *Favours screening/control or no difference* • *Substantial /moderate/mild heterogeneity*[a] P=X (I^2=X)
Localised (T1-T2, N0, M0) PLCO ERSPC Norrkoping	3 (247,954)	10,107/112,725 (8.97)	7671/135,229 (5.67)	1.79 (1.19-2.70)	• Favours control P=0.005 • Substantial heterogeneity P<0.00001 (I^2=99%)
	NB. Incorporating data from the French ERSPC site resulted in no change in findings (RR 1.66; 95% CI 1.22-1.27)				
Advanced (T3-T4, N1, M1) PLCO ERSPC Norrkoping	3 (247,954)	868/112,725 (0.77)	1460/135,229 (1.08)	0.80 (0.73-0.87)	• Favours screening P<0.00001 • No significant heterogeneity P=0.51 (I^2=0%)
	NB. Incorporating data from the French ERSPC site resulted in no change in findings (RR 0.77; 95% CI 0.71-0.83)				
Authors' comments: • The quality of evidence was rated as LOW for localised prostate cancer and MODERATE for advanced prostate cancer according to the GRADE approach • The quality of the individual RCTs was assessed as: - The ERSPC and PLCO studies were assessed as a low risk of bias - The Norrkoping study was graded as a high risk of bias • Sensitivity analysis demonstrated that the exclusion of the Norrkoping study resulted in a reduction in the effectiveness of screening in detecting localised prostate cancer (RR 1.44; 95% CI 0.90-2.32) but no effect on advanced cancer					
Other results					
Prostate cancer-specific metastatic disease	• "There were very limited data on metastatic disease." • See the results for "Advanced prostate tumour stage" above				
Skeletal-related events	NR				
Quality of life	• "None of the studies provided a complete assessment of the effect of screening on quality of life." • Both the ERSPC and PLCO studies are currently assessing measures relating to quality of life				
Test performance characteristics	• False-positive rate for men who had an elevated PSA value was 17.8% for men screened at least once in the ERSPC study compared to a detection rate of 3.4-3.6%. • The rate of overdiagnosis in the screening group was estimated to be up to 50% in the ERSPC study • The PLCO study reported a false-positive rate of 10.4% for the PSA test and 15.0% for DRE				
Harms and benefits of screening	• Common minor harms from screening include bleeding, bruising and short term anxiety • Common major harms include overdiagnosis and overtreatment, infection, blood loss requiring transfusion, pneumonia, erectile dysfunction and incontinence • The ongoing CAP study reports on a variety of harms associated with screening (Rosario 2012). These include mild or no pain, dizziness, fever, haematuria, haematochezia and haemoejaculate				
Harms and benefits of PSA test	• PLCO study reported on the complications rate of the PSA test. Complications included dizziness, bruising, hematoma and fainting				
Harms and benefits of biopsy	• PLCO study reported on the complications from diagnostic procedures after a positive PSA test result. Complications included primary infection, bleeding, clot formation and urinary difficulties • ERSPC study reported on the complications from the biopsy procedure • The Netherlands site of ERSPC (Raaijmakers et al, 2002) reported on the complications rate of TRUS guided sextant biopsies. Complications included hematospermia and hematuria for greater than 3 days. The most common side effects were pain after biopsy and fever				
Harms and benefits of treatment	• PLCO study reported on adverse events from treatment				

Authors' comments	
Data from the PLCO trial	The PLCO study reports on 10 and 13 year follow-up of participants. However, the 10 year data is complete for 92% of participants, whilst the 13 year data is complete for only 57% of participants. Consequently, the Cochrane review used the 10 year follow-up data for all of their analyses with the exception of the analysis on tumour stage, which used the 13 year follow-up data

Bias	Authors' quality assessment of included studies • *ERSPC:* low risk of bias • *Norrkoping:* high risk of bias due to high risk associated with the allocation sequence generation and allocation concealment as well as uncertainty about incomplete outcome data • *PLCO:* low risk of bias • *Quebec:* high risk of bias due to high risk associated with allocation concealment and analysing data not using the intention-to-treat principle, as well as uncertainty about random sequence generation, blinding of outcome assessors and selective reporting • *Stockholm:* high risk of bias due to high risk associated with allocation concealment and uncertainty with sequence generation. This study also had low external validity as it had a one-time screen for prostate cancer, with biopsy only performed if PSA >10 ng/mL Other potential sources of bias • *ESRPC study* - It was decided that an age range of 55-69 years would be the 'core' age group for participants. The inclusion of higher and/or lower age groups was left to the discretion of the participating centres - The primary endpoint was the rate of prostate cancer mortality in the total study arm compared with the control arm, with one analysis to be conducted for the 'core' age group and another for all ages at entry - ERSPC had sufficient power to detect a significant difference in prostate cancer mortality between the total study arm compared with the control arm if the true reduction in mortality by screening was 25% or more, or if contamination was limited to 10% if the true effect was 20% or more - The screening protocol was changed during the ERSPC study. Both the DRE and TRUS ceased to be used as screening tests in 1997. The PSA cut-off value was also reduced • *Quebec study:* data were not analysed according to the intention-to-treat principle • Information could not be obtained for the following variables to assess bias: - How allocation concealment was obtained in the ERSPC study - Quebec and Stockholm studies provided insufficient information to determine how sequence generation was performed - Quebec study did not provide clear information about how blinding of outcome assessment was achieved - Withdrawals were cited in the Norrkoping study but it was unclear how data for men who participated but migrated out of the catchment area were obtained
Contamination	• The contamination rate in the ERSPC study was estimated to be 30.7%, accounting for 27,431/89,353 men in the control group having at least one PSA test (Roobol et al, 2009) • PLCO reported that 45% of participants entered the study with a history of PSA testing in the three years prior to randomisation. 52% of men assigned to the control group underwent some form of screening during the study period • There was potential for contamination in the Norrkoping trial as study details were distributed through the newspaper, radio and television advertisements • In the Quebec trial, only 23.6% of participants randomised to the screening group were screened. 7.3% of participants in the control group were screened
Internal validity	
Overall quality rating	Good
Comments	None
External validity	
Generalisability	The Cochrane review included all men enrolled in studies of prostate cancer screening with no exclusions based on ethnicity, age or presence of LUTS. Men with a previous diagnosis and treatment of prostate cancer were excluded. The age range of participants within the included studies was between 45-80 years

Applicability	None of the studies included in the Cochrane review were conducted in Australia. The PSA test cut-off was not specifically defined and was dependent on the PSA cut-off used in each individual study
Comments - Contamination?	See section on "Authors' comments – contamination"
References	
Raaijmakers R, Kirkels W, Roobol M, Wildhagen M, and Schroder F. (2002). Complication rates and risk factors of 5802 transrectal ultrasound-guided sextant biopsies of the prostate within a population-based screening program. *Urology* 60:826–30. Roobol M, Kerkhoff M, Schroder F, Cuzik J, Sasieni P, Hakama M, et al. (2009). Prostate cancer mortality reduction by prostate-specific antigen based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 56:585–91.	

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; GRADE, Grading of Recommendations Assessment, Development and Evaluation; LUTS, lower urinary tract symptoms; MA, meta-analysis; NHS EED, National Health Service Economic Evaluation Database; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; SR, systematic review; TRUS, transrectal ultrasound.

[a] Heterogeneity defined as follows: (i) no significant heterogeneity if Phet>0.1 and I^2<25%; (ii) mild heterogeneity if I^2 <25%; moderate heterogeneity if I^2 between 25-50%; substantial heterogeneity I^2 >50%.

Study ID	**Lumen et al (2012)**
Citation/Primary publication	Lumen N, Fonteyne V, de Meerleert G, Ost P, Villeirs G, Mottrie A, de Visschere P, de Troyer B, and Oosterlinck, W. (2012). Population screening for prostate cancer: an overview of available studies and meta-analysis. *Int J Urol* 19:100-108.
Study type (Level)	SR MA (I) The authors included the Rotterdam-Ireland trial which it notes was not a prospective RCT. Rather, it was a comparison between a screened population (part of the Rotterdam section of the ERSPC trial) and a population where screening is not routinely carried out (Ireland). Consequently, the evidence reviewer acknowledges that Lumen (2012) does fit precisely into NHMRC's classification of a Level I study.
Affiliation/Source of funds	Ghent University Hospital, Belgium
Intervention	
Description of test +/- DRE +/- TRUS	PSA test with or without DRE
PSA test cut-off	Not specifically defined – dependent on the PSA test cut-off used in each individual study
Comparator	
Description of comparator	No screening for prostate cancer
Eligibility criteria	
Inclusion criteria	Asymptomatic men
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	None reported
Literature search	
Search strategy Search period? Publication types? English only?	• *Search period*: last search date 4 April 2011 • *Publication types:* systematic reviews, meta-analyses, RCTs and comparative trials dealing with screening versus no screening for prostate cancer • Eligible studies were included if they provided personalised data on one or more of the following endpoints: prostate cancer incidence, prostate cancer stage or grade at diagnosis, prostate cancer mortality or overall mortality
Exclusion criteria	• The Medline search was limited to: humans, gender (male), language (English) and article type (RCT or comparative study) • No limits were applied to the Web of Science database search
Databases searched	Medline (PubMed) and Web of Science databases
Outcomes and measures	
Primary outcome	• Prostate cancer incidence • Prostate cancer stage or grade at diagnosis • Prostate cancer mortality • Overall mortality
Secondary outcome	None
Other outcomes	None

Key results					
Outcome	**No. trials (no. patients)**	**Screening n/N (%)**	**No screening n/N (%)**	**Relative risk estimate (95% CI)**	**P-value** • *Favours screening/no screening or no difference* • *Substantial /moderate/mild heterogeneity*[a] P=X (I^2=X)
Diagnosis of prostate cancer					
Incidence of prostate cancer Norrkoping Stockholm French ERSPC Goteborg Rotterdam-Ireland PLCO ERSPC	7 (525,108)	12,447/179,639 (6.93)	14,413/345,469 (4.17)	1.55 (1.17-2.06)	• Favours screening P=0.002 • Substantial heterogeneity P<0.00001 (I^2=99%)
Localised prostate cancer Norrkoping French ERSPC PLCO ERSPC	4 (332,743)	8832/155,317 (5.69)	5850/177,426 (3.30)	1.81 (1.15-2.86)	• Favours screening P=0.01 • Substantial heterogeneity P<0.00001 (I^2=99%)
Metastatic disease French ERSPC Rotterdam-Ireland Norrkoping Goteborg PLCO ERSPC	6 (497,945)	281/177,259 (0.16)	1360/320,686 (0.42)	0.63 (0.38-1.05)	• No difference P=0.079 • Substantial heterogeneity P<0.00001 (I^2=88%)
Low-grade prostate cancer (Gleason ≤6) Norrkoping Goteborg French ERSPC Rotterdam-Ireland PLCO ERSPC	6 (497,945)	7682/177,259 (4.33)	5601/320,686 (1.75)	2.32 (1.39-3.88)	• Favours screening P=0.001 • Substantial heterogeneity P<0.00001 (I^2=99%)
High-grade prostate cancer (Gleason ≥8) Norrkoping Goteborg French ERSPC Rotterdam-Ireland PLCO ERSPC	6 (497,945)	820/177,259 (0.46)	1863/320,686 (0.58)	0.91 (0.73-1.14)	• No difference P=0.42 • Substantial heterogeneity P=0.0008 (I^2=76%)

Prostate cancer-specific mortality					
Prostate cancer-specific mortality PLCO Norrkoping Goteborg Rotterdam-Ireland Stockholm Quebec ERSPC	7 (486,813)	579/168,182 (0.34)	1786/318,631 (0.56)	0.88 (0.72-1.06)	• No difference P=0.18 • Substantial heterogeneity P=0.009 (I^2=65%)
Prostate cancer-specific mortality[b] Norrkoping Goteborg Rotterdam-Ireland ERSPC	4 (336,430)	323/96,306 (0.34)	1161/240,124 (0.48)	0.76 (0.58-0.98)	• Favours screening P=0.04 • Substantial heterogeneity P=0.03 (I^2=66%)
Overall (all-cause) mortality					
Overall mortality Goteborg Stockholm Rotterdam-Ireland PLCO	4 (269,058)	8596/62,665 (13.72)	43,451/206,393 (21.05)	0.90 (0.75-1.08)	• No difference P=0.27 • Substantial heterogeneity P<0.00001 (I^2=98%)
Overall mortality[b] Goteborg Rotterdam-Ireland	2 (165,161)	3657/21,922 (16.68)	29,065/143,239 (20.29)	0.83 (0.58-1.20)	• No difference P=0.32 • Substantial heterogeneity P<0.00001 (I^2=99%)
Other results					
Prostate cancer-specific metastatic disease	NR				
Skeletal-related events	NR				
Quality of life	NR				
Test performance characteristics	NR				
Harms and benefits of biopsy	NR				
Harms and benefits of treatment	NR				
Other harms and benefits	NR				

Authors' comments	
Main shortcomings of the individual studies	• Norrkoping - During the first two screening rounds, only DRE was used as a screening tool - Fine needle aspiration was used during prostate biopsy which has a lower accuracy compared with true cut biopsy - A low number of patients in the screening group led to a low number of prostate cancer in this group - 48 localised and thus curable tumours were diagnosed in the screening group but only 21 were treated with curative intention • Quebec - Participation in the screening group was only 24% - The only data available was on prostate cancer mortality • ERSPC - Substantial heterogeneity in screening interval, screening strategy and PSA threshold in the different centres of the study - PSA contamination in the non-screening group was estimated to be 30.9% • PLCO - Short median follow-up of 7 years - High rate of PSA contamination (52%) in the non-screening group - In both groups, PSA testing was carried out in 44% of included men within the year before randomisation - Compliance rate to undergo prostate biopsy in the case of a positive screening test was only 30-40% • Goteborg - Relatively low number of patients in both groups - Some of the patients were already included in the ERSPC trial • Rotterdam-Ireland - Not a prospective randomised clinical trial but a comparison between a screened population and a population where screening is not routinely carried out • French ERSPC - Follow-up of just 4 years - Low participation rate in the screening group (27.7%) - Low compliance (45.9%) for subsequent prostate biopsy in the case of a positive screening test in the screening group • Stockholm - Screening included only one PSA test with DRE and TRUS - High PSA threshold (10 ng/mL) - Low number of patients in the screening group - Insufficient treatment of curable prostate cancer - Reconstruction of the no screening group
Internal validity	
Overall quality rating	Good
Comments	None
External validity	
Generalisability	The systematic review by Lumen et al (2012) included asymptomatic men. No exclusion criteria were applied. The age range of participants within the included studies was between 45-80 years
Applicability	None of the studies included in Lumen et al (2012) were conducted in Australia. The PSA test cut-off was not specifically defined and was dependent on the PSA cut-off used in each individual study
Comments - Contamination?	See section on "Authors' comments"

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; MA, meta-analysis; NHMRC, National Health and Medical Research Council; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; SR, systematic review; TRUS, transrectal ultrasound.

[a] Heterogeneity defined as follows: (i) no significant heterogeneity if Phet>0.1 and I^2<25%; (ii) mild heterogeneity if I^2 <25%; moderate heterogeneity if I^2 between 25-50%; substantial heterogeneity I^2 >50%.

[b] Adjusted analysis after the exclusion of studies with follow-up <8 years (PLCO, French ERSPC), PSA contamination in the non-screening group >33.3% (PLCO) and participation in the screening group <75% (Quebec, Stockholm and French ERSPC). These criteria were defined by the authors.

Study ID	**Djulbegovic et al (2010)**
Citation/Primary publication	Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, and Dahm P. (2010). Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. 2010. *BMJ* 341:c4543.
Study type (Level)	SR MA (I)
Affiliation/Source of funds	• Department of Urology, University of Florida • Dennis W Jahnigen Career Development Scholars Award through the American Geriatrics Society *"The funders had no role in study design, the collection, analysis and interpretation of data, the writing of the report or the decision to submit the article for publication."*
Intervention	
Description of test +/- DRE +/- TRUS	PSA test with or without DRE
PSA test cut-off	Not specifically defined – dependent on the PSA test cut-off used in each individual study
Comparator	
Description of comparator	No screening for prostate cancer (by PSA test with or without DRE)
Eligibility criteria	
Inclusion criteria	Screening of asymptomatic men for prostate cancer
Definition of *asymptomatic*	Not specifically defined. But men without a previous history of prostate cancer are mentioned
Exclusion criteria	Men with a previous diagnosis of prostate cancer
Literature search	
Search strategy Search period? Publication types? English only?	• *Search period:* 1 January 2005 to 13 July 2010 (this was an updated search of the Cochrane review published in 2006) • *Publication types:* RCTs
Inclusion criteria	• RCTs comparing screening of asymptomatic men by PSA testing with or without DRE versus no screening • Eligible studies were included irrespective of language or publication status
Exclusion criteria	Trials with participants with previously diagnosed prostate cancer
Databases searched	• Medline (PubMed), EMBASE, CENTRAL (Cochrane Central Register of Controlled Trials) • Manual search of abstract proceedings from the American Urological Association (AUA), European Association of Urology (EAU) and American Society of Clinical Oncology (ASCO) meetings from 2005-2010 • Search of additional systematic reviews and narrative reviews on the topic to identify eligible trials
Outcomes and measures	
Primary outcome	All-cause mortality and death from prostate cancer
Secondary outcome	• Diagnosis of prostate cancer • Effect of screening on stage at diagnosis • False-positive and false-negative results • Harms of screening • Quality of life • Cost effectiveness
Other outcomes	None

Key results					
Outcome	**No. trials (no. patients)**	**Event rate (per 1000) with screening (95% CI)**	**Event rate (per 1000) with control (95% CI)**	**Relative risk estimate (95% CI)**	**P-value** • *Favours screening/ control or no difference* • *Substantial /moderate/mild heterogeneity*[a] P=X (I^2=X)
All-cause mortality[b] ERSPC 2009 Gothenburg 2010 Norrkoping 2004 PLCO 2009 *Quality of evidence (GRADE): Moderate*	4 (256,019)	198 (194-202)	200 (NR)	0.99 (0.97-1.01)	• No difference P=0.44 • No significant heterogeneity P=0.60 (I^2=0%, X^2 =1.89)
Deaths from prostate cancer[b] ERSPC 2009 Gothenburg 2010 Norrkoping 2004 PLCO 2009 Quebec 2004 *Quality of evidence (GRADE): Moderate*	5 (302,500)	7 (6-9)	8 (NR)	0.88 (0.71-1.09)	• No difference P=0.25 • Substantial heterogeneity P=0.06 (I^2=55%, X^2 =8.89)
Outcome	**No. trials (no. patients)**	**Screening n/N (%)**	**Control n/N (%)**	**Relative risk estimate (95% CI)**	**P-value** • *Favours screening/ control or no difference* • *Substantial /moderate/mild heterogeneity*[a] P=X (I^2=X)
Diagnosis of prostate cancer					
Total PLCO 2009 French ERSPC 2009 Norrkoping 2004 ERSPC 2009 Gothenburg 2010 *Quality of evidence (GRADE): Low*	5 (340,800)	10,328/159,372 (6.48)	7968/181,428 (4.39)	1.46 (1.21-1.77)	• Favours screening P<0.001 • Substantial heterogeneity P<0.001 (I^2=97%, X^2 =126.69)
Subgroup analysis: stage of prostate cancer[c]					
Stage I ERSPC 2009 French ERSPC 2009 Norrkoping 2004 PLCO 2009 *Quality of evidence (GRADE): Low*	4 (332,743)	3789/155,317 (2.44)	1971/177,426 (1.11)	1.95 (1.22-3.13)	• Favours screening P=0.005 • Substantial heterogeneity P<0.001 (I^2=96%, X^2 =79.32)

Stage II ERSPC 2009 French ERSPC 2009 Norrkoping 2004 PLCO 2009 *Quality of evidence (GRADE): Very low*	4 (332,743)	5114/155,317 (3.29)	4035/177,426 (2.27)	1.39 (0.99-1.95)	• No difference P=0.05 • Substantial heterogeneity P<0.001 (I^2=97%, X^2=114.38)
Stage III and IV ERSPC 2009 French ERSPC 2009 Norrkoping 2004 PLCO 2009 *Quality of evidence (GRADE): Moderate*	4 (332,743)	701/155,317 (0.45)	975/177,426 (0.55)	0.94 (0.85-1.04)	• No difference P=0.22 • No significant heterogeneity P=0.75 (I^2=0%, X^2=1.22)
Prostate cancer-specific metastatic disease	See above for diagnosis of Stage III and IV prostate cancer				
Skeletal-related events	NR				
Quality of life	"No study reported data on quality of life. Though the ERSPC study collected data on quality of life, no detailed analyses have been made available to date."				
Other results					
Subgroup analysis based on age	• *All-cause mortality:* age specific information limited to and reported in the Gothenburg study • *The effect of screening on death from prostate cancer*: age specific information limited to ERSPC with additional information from the Goteborg study only for men aged 50-54 years • *Age specific diagnosis of prostate cancer based on DRE alone*: reported for the first two screening rounds of Norrkoping. Goteborg contributed data based on PSA testing for men aged 50-54, 55-59 and 60-64				
Test performance characteristics	• False-positive rates of screening are reported in the Norrkoping (82.5%) and ERSPC (75.9%) studies				
Harms and benefits of biopsy	• A recent abstract (Carlsson et al, 2010) based on three ERSPC study centres reported no excess mortality associated with prostate biopsies in the screening arm				
Harms and benefits of treatment	NR				
Other harms and benefits	• PLCO study reported on the complications of PSA test and a DRE. Complications included bleeding, pain, fainting, infections, clot formation and urinary difficulties • The ongoing ProtecT study in the UK and its extension, the comparison arm for ProtecT (CAP) trial may provide further evidence on the benefits and harms of screening. Final reports are not expected until 2013 and 2015, respectively				
Authors' comments					
Limitations of included studies	• Potential biases in the individual trials include lack of allocation concealment or intention-to-screen analysis which is expected to favour the screening arm • Contamination of non-screening arm is explicitly reported in the PLCO study but is a possible issue in all studies. This potentially introduces a bias towards not finding a benefit of screening • The short length of follow-up of reported studies may not be enough time to detect differences in mortality given the low number of deaths from prostate cancer • Insufficient evidence to analyse the impact of screening on high risk populations such as patients with a strong family history of prostate cancer or African Americans • Lack of available data to analyse the effect of screening interventions based on participants' age				
Internal validity					
Overall quality rating	Good				
Comments	None				
External validity					
Generalisability	The systematic review by Djulbegovic et al (2010) included all asymptomatic men enrolled in studies of prostate cancer screening with the exclusion of men with a previous diagnosis of prostate cancer. The age range of participants within the included studies was between 45-80 years				

Applicability	None of the studies included in Djulbegovic et al (2010) were conducted in Australia. The PSA test cut-off was not specifically defined and was dependent on the PSA cut-off used in each individual study
Comments - Contamination?	See section on "Authors' comments"
References	
Carlsson SC, Holmberg E, Auvinen AP, Moss SM, Roobol MJ, Schroder FH et al. No excess mortality after prostate biopsy: results from the European randomized study of screening for prostate cancer (ERSPC). European Association of Urology Annual Meeting. Barcelona: European Association of Urology (EAU), April 19, 2010. www.uroweb.org/publications/abstracts-online/?id=108&no_cache=1&AID=26663.	

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; GRADE, Grading of Recommendations Assessment, Development and Evaluation; MA, meta-analysis; NR, not reported; PLCO, Prostage, Lung, Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; SR, systematic review; TRUS, transrectal ultrasound.

[a] Heterogeneity defined as follows: (i) no significant heterogeneity if Phet>0.1 and I^2<25%; (ii) mild heterogeneity if I^2 <25%; moderate heterogeneity if I^2 between 25-50%; substantial heterogeneity I^2 >50%.

[b] The inverse variance method was used as event rates were not available in all studies.

[c] The 2010 American Joint Committee on Cancer system was used for prostate cancer staging.

Study ID	**Lin et al (2011) [AHRQ]**
Citation/Primary publication	Lin K, Croswell JM, Koenig H, Lam C, and Maltz A. (2011). Prostate-specific antigen-based screening for prostate cancer: an evidence update for the U.S. preventive services task force. Evidence synthesis no. 90. AHRQ publication no. 12-05160-EF-1. Rockville, MD: Agency for Healthcare Research and Quality. Associated publication: Chou R, Croswell JM, Dana T, Bougatsos C, Blazina I, Fu R, Gleitsmann K, Koenig HC, Lam C, Maltz A, Rugge JB and Lin K. (2011). Screening for prostate cancer: a review of the evidence for the U.S. preventative services task force. *Ann Intern Med* 155:762-771.
Study type (Level)	SR (I)
Affiliation/Source of funds	Agency for Healthcare Research and Quality
Intervention	
Description of test +/- DRE +/- TRUS	PSA testing (all modalities) with or without DRE and TRUS e.g. single threshold PSA testing as well as other PSA-based prognostic measures such as age-adjusted thresholds, velocity and doubling time
PSA test cut-off	Not specifically defined – dependent on the PSA test cut-off used in each individual study
Comparator	
Description of comparator	No PSA screening or usual care in the asymptomatic general primary care population
Eligibility criteria	
Inclusion criteria	PSA-based screening of asymptomatic men for prostate cancer
Definition of *asymptomatic*	• *PSA-based screening*: a screening program for prostate cancer in asymptomatic men that incorporates one or more PSA measurements, with or without additional modalities such as DRE and TRUS • *Asymptomatic*: without symptoms that are highly suspicious for prostate cancer. Many older men have chronic, stable LUTS (e.g. due to BPH) that are not generally associated with an increased risk for prostate cancer
Exclusion criteria	Not specifically defined
Literature search	
Search strategy Search period? Publication types? English only?	• *Search period:* 1 January 2007 to 1 July 2011 (this was an updated search of the USPSTF review published in 2008) • *Publication types:* RCTs, systematic reviews and meta-analyses only • English language only
Inclusion criteria	Studies that compared PSA-based screening with no screening or usual care in asymptomatic general primary care populations and reported prostate cancer or all-cause mortality as an outcome
Exclusion criteria	• Non-randomised analysis of an RCT, narrative review, editorial or commentary • Articles not in the English language
Databases searched	• PubMed, Cochrane Database • Hand searching of reference lists from included studies and review articles, and recommendations of experts
Outcomes and measures	
Primary outcome	1) Does PSA-based screening decrease prostate cancer-specific or all-cause mortality? 2) What are the harms of PSA-based screening for prostate cancer?
Secondary outcome	None
Other outcomes	None

Key results	
1) Does PSA-based screening decrease prostate cancer-specific or all-cause mortality?	
Included RCTs and their quality rating	• Sandblom et al, 2004; Sandblom et al, 2011 (Norrkoping) – Poor quality • Kjellman et al, 2009 (Stockholm) – Poor quality • Andriole et al, 2009 (PLCO) – Fair quality • Schroder et al, 2009 (ERSPC) – Fair quality • Hugosson et al, 2010 (ERSPC – Goteborg) – Fair quality • Labrie et al, 2004 (Quebec) – Poor quality Two met-analyses were also included: • Djulbegovic et al, 2010 • Ilic et al, 2011
Authors' findings	• "After about 10 years, PSA-based screening results in the detection of more cases of prostate cancer, but small to no reduction in prostate cancer-specific mortality." • "Most RCTs have not reported an effect of PSA-based screening on prostate cancer mortality."
2) What are the harms of PSA-based screening for prostate cancer?	
Included RCTs	• ERSPC • PLCO
Test performance characteristics	• The Finnish centre of the ERSPC trial (Kilpelainen et al, 2010) reported that after three rounds of PSA testing, 12.5% of participants received at least one false-positive result. • In the entire ERSPC trial, 75.9% of men that underwent a biopsy because of an elevated PSA value had a false-positive result • The PLCO trial reported that after four PSA tests, men had a 12.9% cumulative risk of receiving at least one false-positive result and a 5.5% risk of having at least one biopsy as a direct consequence of a false-positive screening test
Harms and benefits of biopsy/diagnostic procedures	• PLCO study reported on the physical harms of screening. These included rare bleeding or pain from DRE. Complications of diagnostic procedures included infection, bleeding and urinary difficulty • ERSPC study noted that no deaths were associated with the prostate biopsy procedure • Rotterdam study centre of ERSPC (Raaijmakers et al, 2002) reported on site-specific biopsy-related harms. These included fever, urinary retention, hospitalisation for signs of prostatitis or urosepsis, hematuria and hematospermia
Quality of life	• "None of the RCTs of PSA-based screening and prostate cancer mortality provided information on potential psychological harms of prostate cancer screening, such as anxiety or impact on health-related quality of life." • The 2008 evidence review performed for the USPSTF found evidence that false-positive PSA test results are associated with adverse psychological effects but could not determine the exact magnitude of psychological harms of prostate cancer screening
Other results	
All-cause mortality	NR
Prostate cancer-specific metastatic disease	NR
Skeletal-related events	NR
Harms and benefits of treatment	NR
Other harms and benefits	NR

Authors' comments	
Authors' quality assessment of included studies	• *Norrkoping*: rated POOR quality due to inadequate method of randomisation; no information available on baseline comparability of the screened and control group; no information on the degree of contamination in the control group; insufficient information regarding outcome assessment; sample size was originally calculated to assess the acceptance and feasibility of a prostate cancer screening program rather than mortality outcomes • *Stockholm*: rated POOR quality due to uncertainty about initial comparability of the screening and comparison groups; potential for attribution bias in outcome assessment; internal discrepancies about the total number of participants • *PLCO*: rated FAIR quality due to high rates of contamination in the control group • *ERSPC*: rated FAIR quality due to inconsistencies in screening intervals and PSA cut-off points among study centres; differences in exclusion of eligible and randomised men by age between centres; exclusion of data from two study centres (Portugal and France); contamination was not evaluated at all centres for the duration of the trial • *Goteborg*: rated FAIR quality due to lack of information regarding baseline comparability of the two arms, attrition and contamination rates
Possible reasons why RCTs have not reported an effect of PSA-based screening on prostate cancer mortality	• Incomplete follow-up and the findings of the two biggest trials (PLCO and ERSPC) may change with additional follow-up • Neither PLCO nor ERSPC excluded men who had a history of PSA testing • Issue of contamination in the control and noncompliance in the intervention arm • Differences in PSA cut-off points, screening intervals and treatment choices
Internal validity	
Overall quality rating	Good
Comments	• Only PubMed and Cochrane Database were searched • No meta-analysis was completed by the authors of the SR. Rather, the results of the individual included studies were discussed and a descriptive overall conclusion was drawn by the authors • Sources of heterogeneity were not explored
External validity	
Generalisability	The systematic review by Lin et al (2011) included all asymptomatic men enrolled in studies of PSA-based screening for prostate cancer. No exclusion criteria were defined. The age range of participants within the included studies was between 45-80 years
Applicability	None of the studies included in Lin et al (2011) were conducted in Australia. The PSA test cut-off was not specifically defined and was dependent on the PSA cut-off used in each individual study
Comments - Contamination?	See section on "Authors' comments'
References	
Kilpeläinen TP, Tammela TL, Määttänen L, Kujala P, Stenman UH, Ala-Opas M, et al. (2010). False-positive screening results in the Finnish prostate cancer screening trial. *Br J Cancer* 102(3):469-74. Raaijmakers R, Kirkels WJ, Roobol MJ, Wildhagen MF, and Schröder FH. (2002). Complication rates and risk factors of 5802 transrectal ultrasound-guided sextant biopsies of the prostate within a population-based screening program. *Urology* 60(5):826-30.	

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; BPH, benign prostatic hyperplasia; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; LUTS, lower urinary tract symptoms; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; SR, systematic review; TRUS, transrectal ultrasound; USPSTF, U.S. Preventative Services Task Force.

Study ID	NZGG (2009)
Citation/Primary publication	New Zealand Guidelines Group. (2009). Cancer control strategy guidance completion: update of evidence for prostate-specific antigen (PSA) testing in asymptomatic men. Wellington: Ministry of Health.
Study type (Level)	SR (I) NB. All levels of evidence were included but Levels I and II evidence were reported separately
Affiliation/Source of funds	Ministry of Health, New Zealand
Intervention	
Description of test +/- DRE +/- TRUS	PSA testing (all modalities) and surveillance versus early treatment
PSA test cut-off	Not specifically defined – dependent on the PSA test cut-off used in each individual study
Comparator	
Description of comparator	No PSA testing
Eligibility criteria	
Inclusion criteria	Asymptomatic men over the age of 45 years (inclusive)
Definition of *asymptomatic*	"People who have no symptoms of prostate cancer"
Exclusion criteria	• Studies that assessed the impact of PSA testing on symptomatic men • New technologies (e.g. early testing) or horizon scanning activities
Literature search	
Search strategy Search period? Publication types? English only?	• *Search period:* 2000-2008; except for research question 7 which excluded studies published before 2003 - Search date: April 2009 • *Publication types:* levels of evidence from RCTs to case series. Therefore, the following were included - Systematic reviews of RCTs - RCTs - Pseudo-RCTs - Comparative studies with concurrent controls: non-randomised experimental trials, cohort studies, case-control studies, interrupted time series with a concurrent control group - Comparative studies without concurrent controls: historical control studies, two or more single arm studies, interrupted time series without a parallel control group - Case series • English only journal articles and other peer reviewed publications
Exclusion criteria	• Literature not in the English language • Editorials, comments, book chapters, articles published in abstract form, animal studies, conference proceedings, correspondence or news items from appraisal in the literature review • Non-systematic reviews • Single case/subject designs • Studies with five or fewer participants in either intervention or comparator arm • Studies relating to biopsy procedure • Studies that assessed the impact of PSA testing on symptomatic men
Databases searched	• MEDLINE, CANCERLIT, PROSTATE, NHS EED, CINAHL, EMBASE, Cochrane Library, National Guideline Clearing House (NGC), U.S. Preventive Services Task Force (USPSTF), Agency for Healthcare Research and Quality (AHRQ), Health Services Technology/Assessment Texts (HSTAT), CMA Infobase – Clinical Practice Guidelines, Scottish Intercollegiate Guidelines Network (SIGN), National Institute for Health and Clinical Excellence (NICE), Guidelines International Network (GIN), Canadian Agency for Drugs and Technologies in Health (CADTH), Turning Practice into Research (TRIP), International Network of Agencies for Health Technology Assessment, Medical Services Advisory Committee, Australia and New Zealand Horizon Scanning Network, New Zealand Health Technology Assessment, National Health and Medical Research Council (NHMRC) • Bibliographies of retrieved publications and recent narrative reviews were examined to identify any additional eligible studies • Hand searching of journals and contacting of authors for unpublished research were not undertaken in this review

Outcomes and measures	
Primary outcomes	Seven research questions were asked: 1) Does PSA testing in asymptomatic men alter prostate cancer-related mortality? 2) Does PSA testing in asymptomatic 'high risk' men alter prostate cancer-related mortality? 3) Does PSA testing in asymptomatic men alter the risk of developing prostate cancer-related metastatic disease? 4) Does surveillance or early treatment in asymptomatic men who have a positive PSA test alter prostate cancer-related morbidity or mortality? 5) Of the available modalities of PSA for screening for prostate cancer, what is the sensitivity and specificity of each? 6) Does annual PSA testing versus 3-yearly testing alter prostate cancer-related outcomes in asymptomatic men? 7) What is the incidence of treatment-related mortality and morbidity for prostate cancer in asymptomatic men who had a PSA test?
Secondary outcome	None
Other outcomes	Mortality, morbidity (metastatic disease), sexual dysfunction, urinary incontinence, bowel dysfunction, depression, anxiety
Key results	
1) Does PSA testing in asymptomatic men alter prostate cancer-related mortality?	
Included RCTs and their quality rating	1) Andriole et al, 2009 (PLCO) – good quality 2) Sandblom et al, 2004 (Norrkoping) – mixed quality 3) Schroder et al, 2009 (ERSPC, total) – poor quality 4) Berenguer et al, 2003 (ERSPC, Spain) – mixed quality 5) Nelen et al, 2003 (ERSPC, Antwerp) – mixed quality 6) Labrie et al, 2004 (Quebec) – poor quality
Authors' findings	• Evidence from RCTs is inconsistent and conflicting • Based on available evidence from RCTs, the best estimates for prostate cancer survival at approximately 10 years are 97.2% in the screening population and 95.9% in the control population • Currently, there is no evidence to support or refute a decrease in mortality due to PSA screening. The best case scenario is that there may be a small benefit in survival to men who have been screened
2) Does PSA testing in asymptomatic 'high risk' men alter prostate cancer-related mortality?	
Authors' findings	• *'High risk'* include those with a family history of prostate cancer; *BRCA* positive gene; 70 years and over • No systematic reviews or RCTs were identified which reported on this outcome
3) Does PSA testing in asymptomatic men alter the risk of developing prostate cancer-related metastatic disease?	
Included RCTs and their quality rating	1) Andriole et al, 2009 (PLCO) – good quality 2) Nelen et al, 2003 (ERSPC, Belgium) – good quality 3) Aus et al, 2007 (ERSPC, Sweden) – good quality 4) Berenguer et al, 2003 (ERSPC, Spain) – mixed quality 5) Hugosson et al, 2003; Hugosson et al, 2003b (ERSPC, Sweden) – mixed quality 6) Postma et al, 2006 (ERSPC, Rotterdam) – mixed quality 7) Van der Cruijsen-Koeter et al, 2005 (ERSPC, Rotterdam) – mixed quality 8) Schroder et al, 2009 (ERSPC, total) – poor quality 9) Candas et al, 2000 (Quebec) – poor quality 10) Hugosson et al, 2004 (ERSPC, Sweden) – poor quality 11) Isola et al, 2001 (ERSPC, Finland) – poor quality
Authors' findings	• The presentation of metastatic disease as an outcome varied considerably between RCTs. The data were often descriptive and lacked statistical analysis • The evidence from three RCTs that presented statistical analysis (Schroder et al, 2009; Aus et al, 2007; van der Cruijsen-Koeter, 2005) suggests that the incidence of metastatic disease is reduced in men with a screening program compared with controls • Metastatic disease in screened men is relatively low and early detection and early treatment is likely to further reduce the development of metastatic disease
4) Does surveillance or early treatment in asymptomatic men who have a positive PSA test alter prostate cancer-related morbidity or mortality?	
Surveillance	

Included RCTs and their quality rating	1) Bill-Axelson et al, 2005 – good quality
Authors' findings	• Prostatectomy resulted in a reduction in all mortality endpoints investigated, with a relative reduction of 44% in prostate-related mortality and 26% in overall mortality compared with watchful waiting over a 10 year follow-up period • Prostatectomy also resulted in a reduction of 40% in risk (RR 0.60; 95% CI: 0.42-0.86) of distant metastases compared with watchful waiting over a 10-year follow-up period (absolute risk 10.2; 95% CI: 3.1-17.2; P<0.004)
Early treatment	
Included RCTs and their quality rating	1) Holmberg et al, 2002 – good quality 2) Bill-Axelson et al, 2005 – good quality
Authors' findings	The included studies reported on metastatic disease following prostatectomy. The incidence of metastatic disease following this treatment is minimal in men with early localised prostate cancer treated surgically
5) Of the available modalities of PSA for screening for prostate cancer, what is the sensitivity and specificity of each?	
PSA velocity	
Authors' findings	• *PSA velocity*: rate of increase in PSA levels in the blood as a function of time • No RCTs were identified which reported on this outcome
f/t PSA ratio	
Authors' findings	No RCTs were identified which reported on this outcome
The total approach (absolute level)	
Included RCTs and their quality rating	1) Van der Cruijsen-Koeter et al, 2003 – good quality 2) Auvinen et al, 2004 – poor quality
Authors' findings	Results from the individual RCTs were reported but no conclusions were made by the authors
Overall author comments	Estimates for sensitivity and specificity varied widely, according to total PSA cut-off, study quality and likelihood of bias and age of participants. Additionally verification bias, due to not all participants undergoing biopsies as gold standard testing, is likely to lead to overestimation of true sensitivity and specificity
6) Does annual PSA testing versus 3-yearly testing alter prostate cancer-related outcomes in asymptomatic men?	
Annual PSA testing versus 3-yearly testing	
Authors' findings	No studies were identified which compared annual PSA testing with 3-yearly testing to identify changes in rates of overdiagnosis of prostate cancer
Annual PSA testing versus other interval testing	
Included RCTs and their quality rating	1) Andriole et al, 2009 (PLCO) – good quality (annual screening) 2) Sandblom et al, 2004 (Norrkoping) – mixed quality (3-yearly screening) 3) Nelen et al, 2003 (ERSPC, Belgium) – mixed quality (6-yearly screening) 4) Berenguer et al, 2003 (ERSPC, Spain) – mixed quality (4-yearly screening) 5) Schroder et al, 2009 (ERSPC, total) – poor quality (4-yearly screening) 6) Labrie et al, 2004 (Quebec) – poor quality (annual screening)
Authors' findings	• There was no evidence to support one screening interval over another. • The RCTs showed that annual screenings showed no benefit in reduction in mortality between screened men and controls.
7) What is the incidence of treatment-related mortality and morbidity for prostate cancer in asymptomatic men who had a PSA test?	
Included RCTs and their quality rating	1) Bill-Axelson et al, 2005 – good quality
Authors' findings	It is difficult to reach any conclusions based on the limited data available for this question
Other results	
All-cause mortality	NR
Skeletal-related events	NR
Quality of life	NR
Harms and benefits of biopsy	NR

Harms and benefits of treatment	• Incidence of treatment-related sexual dysfunction is reported in - Systematic review: Mambourg et al, 2006 - Case-control study: Hoffman et al, 2004 - Cohort design: Korfage et al, 2005; Varkarakis et al, 2004; Hoffman et al, 2003; Hoffman et al, 2006; Penson et al, 2005; Penson et al, 2003 • Incidence of treatment-related urinary incontinence is reported in - Systematic review: Mambourg et al, 2006 - Case-control study: Hoffman et al, 2004 - Cohort design: Potosky et al, 2004; Hoffman et al, 2006; Hoffman et al, 2003; Penson et al, 2003; Penson et al, 2005; Korfage et al, 2005; Kwiatkowski et al, 2004; Kwiatkowski et al, 2003 • Incidence of treatment-related bowel dysfunction is reported in - Systematic review: Mambourg et al, 2006 - Case-control study: Hoffman et al, 2004 - Cohort design: Hoffman et al, 2003; Hoffman et al, 2006; Potosky et al, 2004 • Incidence of treatment-related anxiety/depression is reported in Korfage et al, 2005 and Korfage et al, 2006
Other harms and benefits	NR
Internal validity	
Overall quality rating	Good
Comments	• No meta-analysis was completed by the authors of the SR. Rather, the results of the individual included studies were reported and a narrative overall conclusion was drawn by the authors • No tests for heterogeneity were made by the authors
External validity	
Generalisability	The systematic review by NZGG (2009) included asymptomatic men over the age of 45 years
Applicability	None of the studies included in NZGG (2009) were conducted in Australia. The PSA test cut-off was not specifically defined and was dependent on the PSA cut-off used in each individual study
Comments - Contamination?	Not reported

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; NR, not reported; PLCO, Prostate, Lung, Colorectal, Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; RR, relative risk; SR, systematic review; TRUS, transrectal ultrasound.

1.8.2 Randomised controlled trials

Study ID	PLCO
Citation/Primary publication	Andriole GL, Crawford DE, Grubb III RL, Buys SS, Chia D, Church TR, Fouad MN, Isaacs C et al. (2012). Prostate cancer screening in the randomised prostate, lung, colorectal, and ovarian cancer screening trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst* 104:125-132.
Affiliation/Source of funds	National Cancer Institute at the National Institutes of Health
Study characteristics	
Study design	• RCT across 10 study centres in the USA from 1993-2001 – Birmingham AL, Denver CO, Detroit MI, Honolulu HI, Marshfield WI, Minneapolis MN, Pittsburgh PA, Salt Lake City UT, St Louis MO and Washington DC • Each study centre used recruitment sources and strategies appropriate to the local situation. However, all studies enrolled men aged 55-74 years • This study reports on a 13-year follow-up of participants
Randomisation method	• Participants were randomised 1:1 to either the screening or control arms • The randomisation method involved blocks of random permutations of varying lengths, stratified by study centre, gender and age. Random assignment was implemented using compiled software and encrypted files loaded on study centre microcomputers (Prorok et al, 2000)
Location (country)	USA
Inclusion criteria	Men aged 55-74 years who reported no previous personal history of prostate, lung or colorectal cancer
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	• Men who are currently receiving treatment for cancer • Men who have had previous surgical removal of the entire prostate • Men who have previously participated in another cancer screening or primary prevention study • Men who have used finasteride in the previous 6 months • From April 1995: men who had undertaken more than one PSA blood test in the previous 3 years • From April 1995: men who had any lower gastrointestinal diagnostic procedure in the previous 3 years
Study size	• No. eligible men aged 55-70: 76,685 - 38,340 allocated to the intervention arm - 38,345 allocated to the control arm
Length of follow-up *Mean/median (years)*	• Data taken from Andriole et al, 2009 - Median (range): 11.5 (7.2-14.8)
Population	
Age	• 55-74 years
Risk factors	NR
Comorbidities	• A modified Charlson comorbidity score was calculated for eligible participants • The modified Charlson score contained the following conditions found in the Charlson score: myocardial infarction, stroke, diabetes, cancer, pulmonary disease (bronchitis and/or emphysema) and liver disease (cirrhosis and/or hepatitis) • Not included in the modified score were congestive heart failure, peripheral vascular disease, connective tissue disease, hemiplegia, HIV, renal disease, ulcer disease and dementia • A score of 0 = no comorbidity; ≥1 = one or more comorbid conditions

Other key population characteristics	N/A			
Intervention				
Description of test +/- DRE +/- TRUS	• Participants in the screening group were offered annual PSA testing for 6 years and annual DRE for 4 years • A positive test was defined as PSA >4.0ng/mL or a suspicious DRE			
PSA test cut-off	4.0 ng/mL			
Frequency	Annually			
Comparator				
Description of comparator	Usual care, where men sometimes underwent opportunistic screening when a test was requested by a participant or recommended by a doctor			
Outcomes				
Primary outcome	Prostate cancer-specific mortality at 13 years follow-up			
Secondary outcome	All-cause mortality, incidence of prostate cancer, staging and survival			
Other outcomes	N/A			
Key results				
Population analysed	**Intervention**		**Comparator**	
Randomised	38,340		38,345	
Efficacy analysis (ITT)	38,340		38,345	
Efficacy analysis (PP)	NR		NR	
Safety analysis	NR		NR	
Outcome	**Screening n/N (%)**	**Control n/N (%)**	**Risk estimate (95% CI)**	**P-value** *Favours screening/control*
Incidence of prostate cancer				
7 years follow-up (Andriole et al, 2009)	2820/NR	2322/NR	1.22 (1.16-1.29)	No difference
10 years follow-up (Andriole et al, 2009)	3452/NR	2974/NR	1.17 (1.11-1.22)	No difference
13 years follow-up	4250/38,340 (11.09)	3815/38,345 (9.95)	1.12 (1.07-1.17)	"Statistically significant 12% increase relative increase in the intervention arm"
Diagnosis of high-grade prostate cancer (Gleason score 8-10) at 13 years of follow-up	401/38,340 (1.05)	454/38,345 (1.18)	0.89 (0.77-1.01)	No difference
Prostate cancer-specific mortality				
Total				
7 years of follow-up (Andriole et al, 2009)	50/NR	44/NR	1.13 (0.75-1.70)	No difference
10 years of follow-up (Andriole et al, 2009)	92/NR	82/NR	1.11 (0.83-1.50)	No difference
13 years of follow-up	158/NR	145/NR	1.09 (0.87-1.36)	No difference

Subgroup analysis: age				
55-64 years	65/NR	54/NR	1.19 (0.83-1.72)	No difference
65-74 years	93/NR	91/NR	1.02 (0.77-1.37)	No difference
Subgroup analysis: comorbidities				
No comorbidities (modified Charlson score of 0)	104/NR	100/NR	1.00 (0.76-1.31)	No difference
With comorbidities (modified Charlson score ≥1)	44/NR	39/NR	1.11 (0.72-1.71)	No difference
Subgroup analysis: pretrial PSA testing				
7 years of follow-up (Andriole et al, 2009)				
≤1 PSA test at baseline	48/34,755 (0.14)	41/34,590 (0.12)	1.16 (0.76-1.76)	No difference
≥2 PSA tests in the previous 3 years at baseline	2/3588 (0.06)	3/3760 (0.08)	0.70 (0.12-4.17)	No difference
10 years of follow-up (Andriole et al, 2009)				
≤1 PSA test at baseline	83/34,755 (0.24)	75/34,590 (0.22)	1.09 (0.80-1.50)	No difference
≥2 PSA tests in the previous 3 years at baseline	9/3588 (0.25)	7/3760 (0.19)	1.34 (0.50-3.59)	No difference
13 years of follow-up				
No pretrial PSA testing	80/NR	64/NR	1.18 (0.85-1.64)	No difference
Any previous pretrial PSA testing	60/NR	59/NR	1.02 (0.71-1.46)	No difference
All-cause mortality (excluding prostate, lung or colorectal cancer)				
7 years of follow-up (Andriole et al, 2009)	2544/NR	2596/NR	0.98 (0.92-1.03)	No difference
10 years of follow-up (Andriole et al, 2009)	3953/NR	4058/NR	0.97 (0.93-1.01)	No difference
13 years of follow-up	5783/NR	5982/NR	0.96 (0.93-1.00)	"Borderline statistical significance"
Prostate cancer-specific metastatic disease	• Stage III incident prostate cancers through 13 years - Intervention: 58/4250 (1.36%) - Control: 65/3815 (1.70%) • Stage IV incident prostate cancers through 13 years - Intervention: 96/4250 (2.26%) - Control: 111/3815 (2.91%)			
Skeletal-related events	NR			

Quality of life	• Quality of life results from the PLCO trial have yet to be reported but are the basis of a future publication • Johnson (2006) reported on the results of a self-administered 36-item health status questionnaire given to participants in the PLCO-Hawaii study site. They found that: - Receiving notification of a cancer diagnosis does not produce an additional negative effect - There is no difference in HRQOL between those in the screened group and controls who were not screened at any of the three time periods; baseline, first follow-up or second follow-up - There was no difference in HRQOL between genders, age groups or ethnicities except that in the screened group there was a reduction in the physical summary score for the oldest age group between base line and first follow-up assessments • Taylor et al (2004) reported on the results of the use of SF-12 to assess the HRQOL of participants in the PLCO-Georgetown University study site. They found: - Participants reported high levels of HRQOL and satisfaction with their decision to participate - Screening arm participants with abnormal screening results had a higher level of intrusive thoughts about cancer than those with all normal results at the short term follow-up but not at the intermediate term follow-up - Trial adherence was statistically significantly better among participants who had received all normal results in the previous year's screening tests than in those who received at least one abnormal result - In the control arm, adherence was positively associated with education and sex
Other results	
Test performance characteristics	NR
Harms and benefits of biopsy	From Andriole et al (2009) • Medical complications from the diagnostic process occurred in 68/10,000 diagnostic evaluations after positive results from screening • Complications were primarily infection, bleeding, clot formation, urinary difficulties
Harms and benefits of treatment	• From Andriole et al (2009) - Treatment-related complications include infection, incontinence, impotence and other disorders - Such complications are being catalogued in a quality of life study and are particularly pertinent in cases of overdiagnosis • Taylor et al (2012) reported on the long-term prostate cancer treatment-related sexual and urinary adverse effects up to 10 years post diagnosis
Other harms and benefits	From Andriole et al (2009) • In the screening group, the complications associated with screening were mild and infrequent • DRE led to a few episodes of bleeding or pain (0.3/10,000 screenings) • PSA led to complications at 26.2/10,000 screenings (primarily dizziness, bruising and hematoma) • PSA led to 3 episodes of fainting per 10,000 screenings
Authors' comments – reasons for the lack of a reduction in mortality	• PSA cut-off of 4 ng/mL and DRE to trigger diagnostic evaluation may not be effective • Contamination in the control could have been substantial enough to dilute any modest effect of annual screening in the screening group • Approximately 45% of men in each study group had undergone one or more PSA tests at baseline, which would have eliminated some cancers detectable on screening from the randomised population • Improvement in therapy for prostate cancer during the course of the trial probably resulted in fewer prostate cancer deaths in the two study groups, which blunted any potential benefits of screening
Internal Validity	

Source of quality assessment	Evidence reviewer
Overall quality rating	Good
Comments	• The vital status of 98% of trial participants was known at 7 years, 92% at 10 years and 57% at 13 years
External validity	
Generalisability	Age range of participants: 55-74 years
Applicability	RCT took place in the USA from 1993-2001. Participants in the screening group were offered annual PSA testing for 6 years and annual DRE for 4 years A positive test was defined as PSA >4.0ng/mL or a suspicious DRE
Comments - Contamination?	• Contamination in control group: increased from 40% in the 1st year to 52% in the 6th year of PSA testing - Approximately 45% of men in each arm had undergone one or more PSA tests prior to trial entry - Using three independently developed models of prostate cancer natural history to conduct a simulated PLCO trial, Gulati et al (2012) found that contamination increased the mortality ratio from 0.68-0.77 to 0.86-0.91, increased the chance of excess mortality in the intervention arm from 0-4% to 15-28%, and decreased the power of the trial to detect a mortality difference from 40-70% to 9-25%. The models indicate that contamination substantially limited the ability of PLCO to identify a clinically significant screening benefit. While the trial shows annual screening does not reduce mortality relative to population screening, contamination prevents concluding whether screening reduces mortality relative to no screening • Compliance rate for screening: 85% for PSA and 86% for DRE • Compliance for prostate biopsy: 30-40%
References	

Andriole GL, Crawford DE, Grub III RL, Buys, SS, Chia D, Church TR, Fouad MN, Gelmann EP et al. (2009) Mortality results from a randomised prostate cancer screening trial. *N Eng J Med* 360:1310-1319.

Gulati R, Tsodikov A, Wever EM, Mariotto AB, Heijnsdijk AM, Katcher J, de Koning HJ, and Etzioni R. (2012). The impact of PLCO control arm contamination on perceived PSA screening efficacy. *Cancer Causes Control* 23:827-835.

Johnson DB (2006). The effects of an abnormal cancer screening test on health related quality of life. *Int J Cancer Res* 2(3): 277-289.

Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, Fogel R, Gelmann EP et al. (2000) Design of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials* 21(6 Suppl): 273S-309S.

Taylor KL, Luta G, Miller AB, Church TR, Kelly SP, Muenz LR, Davis KM, Dawson DL et al. (2012). Long-term disease-specific functioning among prostate cancer survivors and noncancer controls in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *J Clin Oncol* 30(22): 2768-2775.

Taylor KL, Shelby R, Gelmann E, and McGuire C. (2004). Quality of life and trial adherence among participants in the Prostate, lung, Colorectal and Ovarian Cancer Screening Trial. *J Natl Cancer Inst* 96(14): 1083-1094

Abbreviations: AL, Alabama; CI, confidence interval; CO, Colorado; DC, District of Columbia; DRE, digital rectal examination; HI, Hawaii; HRQOL, health-related quality of life; ITT, intention-to-treat; MI, Michigan; MN, Minneapolis; MO, Missouri; N/A, not applicable; NR, not reported; PA, Pennsylvania; PLCO, Prostate, Lung, Colorectal and Ovarian; PP, per-protocol; PSA, prostate-specific antigen; RCT, randomised controlled trial; TRUS, transrectal ultrasound; USA, United States of America; UT, Utah; WI, Wisconsin.

Study ID	**ERSPC**
Citation/Primary publication	Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M et al. (2012). Prostate cancer mortality at 11 years of follow-up. *N Eng J Med* 366:981-990.
Affiliation/Source of funds	• Europe Against Cancer • Fifth and sixth framework program of the European Union • Grants from agencies or health authorities in the participating countries • Unconditional grants from Beckman Coulter • The studies in each national centre were funded by numerous local grants
Study characteristics	
Study design	• RCT across multiple centres in seven European countries – The Netherlands, Belgium Sweden, Finland, Italy, Spain and Switzerland that commenced in 1991 - Portugal was originally part of the trial but discontinued in 2000 as they were unable to provide the necessary data - France was excluded as they commenced participation in 2001 and thus their duration of follow-up was too short (median 4.6 years) • Each country used different recruitment and randomisation procedures. However, all centres included participants in the predefined core age group of 55-69 years who were identified through national registries • This study reports on a 11-year follow-up of participants
Randomisation method	• Within each country, men were assigned to either the screening or control group on the basis of random number generators • Participants were randomised 1:1 in all sites apart from Finland, which took a 2:3 randomisation process • Belgium, The Netherlands, Spain and Switzerland sites were randomised after consent (efficacy trial) • Sweden, Finland and Italy sites were randomised before consent (population-based effectiveness trial)
Location (country)	The Netherlands, Belgium, Sweden, Finland, Italy, Spain and Switzerland
Inclusion criteria	Men aged between 55-74 years
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	Men with an earlier diagnosis of prostate cancer
Study size	• No. eligible men aged 50-74 years: 182,160 - 162,388 men in the predefined core age group of 55-69 years • 82,816 men randomly assigned to the screening group - 72,891 men were 55-69 years • 99,184 men randomly assigned to the control group - 89,352 men were 55-69 years
Length of follow-up *Mean/median (years)*	• Length of follow-up varied between the countries • Overall across all 8 countries (including France) - Median: 9.8 - Mean: 8.6 • Core age group - Median: 11.0 - Mean: 10.5
Population	
Age	• Men aged 50-74 years with the predefined, core age group of 55-69 years - Sweden included men aged 50-54 years - Netherlands, Italy, Belgium and Spain included men up to 74 years at entry - Switzerland included men aged 55-69 years, with screening up to the age of 75 - Finland: men were recruited at the ages of 55, 59, 63 and 67 years and screened until 71 years
Risk factors	NR
Comorbidities	NR
Other key population characteristics	None

Intervention					
Description of test +/- DRE +/- TRUS	Finland	PSA alone			
	Italy	PSA alone			
	Belgium	• PSA + DRE + TRUS (1991-1997) • PSA only (1997 onwards)			
	The Netherlands	• PSA + DRE + TRUS (1991-1997) • PSA only (1997 onwards)			
	Sweden	PSA alone			
	Spain	PSA alone			
	Switzerland	PSA alone			
PSA test cut-off	Finland	• 4.0 ng/mL • Men with PSA 3.0-3.9 ng/mL underwent DRE until 1998			
	Italy	• 4.0 ng/mL • Men with PSA 2.5-3.9 ng/mL underwent DRE and TRUS			
	Belgium	• 10 ng/mL (1991-1994) • 4.0 ng/mL (1995-1997)			
	The Netherlands	• 4.0 ng/mL (1993-1997) • 3.0 ng/mL (1997 onwards)			
	Sweden	• 3.0 ng/mL (1995-1998) • 2.5 ng/mL (1999 onwards)			
	Spain	3.0 ng/mL			
	Switzerland	3.0 ng/mL			
Frequency	• Every 4 years except for Sweden who screened every 2 years • There was a 7-year interval between the 1st and 2nd screening rounds in Belgium • Median screening interval in the core age group: 4.02 years (Schroder et al, 2009)				
Comparator					
Description of comparator	Men not invited for screening				
Outcomes					
Primary outcome	Prostate cancer mortality at 11 years of follow-up				
Secondary outcome	Overall mortality				
Other outcomes	Diagnosis of prostate cancer				
Key results					
Population analysed	**Intervention**		**Comparator**		
Randomised	82,816		99,184		
Efficacy analysis (ITT)	NR		NR		
Efficacy analysis (PP)	NR		NR		
Safety analysis	NR		NR		
Outcome	**Screening n/N (%)**	**Control n/N (%)**	**Risk estimate (95% CI)**	**P-value** *Favours Screening/Control*	
Incidence of prostate cancer among men in the core age group (55-69 years) according to study period					
Total	6963/NR	5396/NR	1.63 (1.57-1.69)	NR	
Study years 1-9	6043/NR	4044/NR	1.88 (1.81-1.96)	NR	
Study years 8-9	1410/NR	1174/NR	1.56 (1.44-1.69)	NR	
Study years 10-11	541/NR	916/NR	0.78 (0.70-0.87)	NR	
Study years 1-11	6584/NR	4960/NR	1.68 (1.62-1.75)	NR	
Study years ≥12	379/NR	436/NR	1.03 (0.9-1.19)	NR	
Prostate cancer-specific mortality among men in the core age group (55-69 years) according to study period					
Subgroup analysis: study period					
Total	299/NR	462/NR	0.79 (0.68-0.91)	Favours screening P=0.001	
Adjusted total (corrected for selection bias and noncompliance)	NR	NR	0.71 (0.58-0.86)	Favours screening P=0.001	
Study years 1-9	189/NR	274/NR	0.85 (0.71-1.03)	No difference P=0.09	

Study years 8-9	71/NR	118/NR	0.74 (0.55-0.99)	Favours screening P=0.04
Study years 10-11	56/NR	111/NR	0.62 (0.45-0.85)	Favours screening P=0.003
Study years 1-11	245/NR	385/NR	0.79 (0.67-0.92)	Favours screening P=0.003
Study years ≥12	54/NR	77/NR	0.80 (0.56-1.13)	No difference P=0.21
Subgroup analysis: age at randomisation				
All ages	364/NR	522/NR	0.83 (0.72-0.94)	Favours screening P=0.005
Core age group	299/NR	462/NR	0.79 (0.68-0.91)	Favours screening P=0.001
≤54 years	6/NR	9/NR	0.65 (0.23-1.83)	No difference
55-59 years	94/NR	144/NR	0.81 (0.62-1.05)	No difference
60-64 years	106/NR	136/NR	0.92 (0.71-1.18)	No difference
65-69 years	99/NR	182/NR	0.67 (0.53-0.86)	Favours screening; P=NR
≥70 years	59/NR	51/NR	1.18 (0.81-1.72)	No difference
All-cause mortality				
Core age group	13,917/NR	17,256/NR	0.99 (0.97-1.01)	No difference P=0.50
All ages	16,737/NR	20,026/NR	1.00 (0.98-1.02)	No difference P=0.85
Other results				
Prostate cancer-specific metastatic disease	• Kerkhof et al (2010) examined the effect of prostate cancer screening on the incidence of metastatic prostate cancer, both with and without adjustment for noncompliance and contamination using data from the Rotterdam ERSPC study site - Prostate cancer screening significantly reduced the occurrence of metastatic prostate cancer [RR 0.75 (0.59-0.95); P=0.02] - Contamination adjusted: RR 0.73 (0.56-0.96); P=0.02 - Noncompliance adjusted: RR 0.72 (0.55-0.95); P=0.02 - Fully adjusted analysis: 0.68 (0.49-0.94); P=0.02 • Schroder et al (2012) examined the effect of prostate cancer screening on the incidence of metastatic disease using data from four centres of the ERSPC (Finland, Netherlands Sweden and Switzerland). Median follow-up of 12 years - Intervention: 256/36,270. Rate of M+ in screening arm: 0.67 (0.59-0.76) - Control: 410/40,543. Rate of M+ in control arm: 0.86 (0.88-1.06) - This translated into a relative reduction of 30% [Hazard Ratio 0.70 (0.60-0.82; P=0.001) in the intention-to-screen analysis and a 42% (P=0.0001) reduction for men who were actually screened			
Skeletal-related events	NR			
Quality of life	• Measures relating to quality of life are currently being reviewed and will form the basis of future publications • Heijnsdijk et al (2012) reports on the predicted number of quality-adjusted life years (QALYs) gained after the introduction of PSA screening using Microsimulation Screening Analysis of ERSPC follow-up data. The model predicted that annual PSA screening of men aged between 55-69 years would result in a gain of 73 life years and 56 QALYs			
Test performance characteristics	• Schroder et al (2009) reported that in the entire ERSPC trial, 75.9% of men that underwent a biopsy because of an elevated PSA value had a false-positive result • The Finnish centre of ERSPC reported that after three rounds of PSA testing (using a cut-off point of 4.0 ng/mL and testing every 4 years), 12.5% of participants received at least one false-positive result • Kilpelainen et al (2011) reported on the number of false-positive screening results from five centres of the ERSPC - Of the 61,604 men who had screened at least once, 17.8% had one or more false-positive			

	result. - Almost 20% of men who participated at all screening rounds had one or more false-positive result - More than half of the men with a false-positive result had another false-positive if screened again
Harms and benefits of biopsy	• No deaths were reported as a direct complication (e.g. septicaemia or bleeding) from the biopsy procedure (Schroder et al, 2009) • An additional study by Carlsson et al (2011) also reported that prostate biopsy is not associated with excess mortality and fatal complications seem to be very rare in a screening setting
Harms and benefits of treatment	NR
Other harms and benefits	NR
Internal Validity	
Source of quality assessment	Evidence reviewer
Overall quality rating	Fair
Comments	None
External validity	
Generalisability	Age range of participants: variable between 50-74 years according to study site but all sites included a predefined core age group of men aged 55-69 years
Applicability	RCT across multiple centres in seven European countries – The Netherlands, Belgium Sweden, Finland, Italy, Spain and Switzerland that commenced in 1991. France is also part of the RCT but its results have not yet been included due to short follow-up. Variable PSA test cut-off according to country (see section on 'Intervention')
Comments - Contamination?	• Contamination in control group: estimated to be 30.7% (Roobol et al, 2009) • Compliance rate for screening: 82.2% • Compliance rate for prostate biopsy: 85.8% • Inconsistencies in screening intervals and PSA cut-off points among study centres • Differences in exclusion of men by age between centres
Additional references	
Carlsson SV, Holmberg E, Moss SM, Roobol MJ, Schroder FH, Tammela TLJ, Aus G, Auvinen AP, and Hugosson J. (2011). No excess mortality after prostate biopsy: results from the European Randomised Study of Screening for Prostate Cancer. *BJU Int* 107:1912-1917. Heijnsdijk EAM, Wever EM, Auvinen A, Hugosson J, Ciatto S, Nelen V, Kwiatkowski M, Villers A et al. (2012). Quality of life effects of prostate specific antigen screening. *N Engl J Med*. 367(7): 595-605. Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schroder FJ, and Steyerberg EW. (2010). Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomised screening trial (ERSPC section Rotterdam). *Int J Cancer* 127:2639-2644. Kilpelainen TP, Tammela TLJ, Roobol M, Hugosson J, Ciatto S, Nelen V, Moss S, Maattaenen L, and Auvinen A. (2011). False-positive screening results in the European randomised study of screening for prostate cancer. *Eur J Cancer* 47:2698-2705. Roobol M, Kerkhoff M, Schroder F, Cuzik J, Sasieni P, Hakama M, et al. (2009). Prostate cancer mortality reduction by prostate-specific antigen based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 56:585–91. Schroder FH, Hugosson J, Carlsson S, Tammela T, Maattanen L, Auvinen A, Kwiatkowski M, Recker F et al. (2012) Screening for prostate cancer decreases the risk of developing metatastic disease: findings from the European Rnadomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 62(5):745-752. Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H et al. (2009) Screening and prostate cancer mortality in a randomized European study. *N Eng J Med* 360:1320-1328.	

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ITT, intention-to-treat; PP, per-protocol; RCT, randomised controlled trial; QALY, quality-adjusted life years; SD, standard deviation; TRUS, transrectal ultrasound.

Study ID	Goteborg
Citation/Primary publication	Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, Pihl CG, Stranne J et al. (2010). Mortality results from the Goteborg randomised population-based prostate cancer screening trial. *Lancet Oncol* 11:725-32.
Affiliation/Source of funds	• The Swedish Cancer Society • The Swedish Research Council • The National Cancer Institute *"The funding sources had no role in the study design and conduct, collection, management, analysis and interpretation of the data, or writing of the report."*
Study characteristics	
Study design	• Ongoing RCT in Goteborg, Sweden that commenced in 1995 • All men aged between 50-64 years as of 31 December 1994 (born 1930-1944) who were living in Goteborg were identified from the population register • 20,000 men were randomly sampled and randomly allocated 1:1 to either a screening group invited for PSA testing or to a control group not invited for screening • This study reports on a 14-year follow-up of participants
Randomisation method	• Sampling and allocation of participants to the intervention or control arm was conducted by computer randomisation • The randomisation procedure was done at the Department of Statistics at the University of Goteborg. • 10-digit personal identifiers were the only available personal data for those doing the computer randomisation
Location (country)	Goteborg, Sweden
Inclusion criteria	Men aged between 50-64 years living in Goteborg, Sweden as of 31 December 1994
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	• Exclusion criteria prior to randomisation but after selection of the sample population - Men with a prior diagnosis of prostate cancer (56 men) - Men who had died (34 men) - Men who had emigrated but had not been removed from the population register at the time of randomisation (six men)
Study size	• No. men born between 1930-1944 (age 50-64 years, median 56 years) living in Goteborg according to the population register: 32,298 • 20,000 men were randomly identified - 19,904 were randomly allocated after excluding ineligible men • 9952 men allocated to the intervention arm (screening group) • 9952 men allocated to the control group
Length of follow-up *Mean/median (years)*	• 14 years • 78% of the randomised men had reached the maximum follow-up time of 14 years (15,501/19,904)
Population	
Age	• 50-64 years • Median: 56 years
Risk factors	NR
Comorbidities	NR
Other key population characteristics	NR
Intervention	
Description of test +/- DRE +/- TRUS	PSA test alone
PSA test cut-off	• 1995-1998: 3.4 ng/mL (WHO corrected value; the nominal value was 3.0 ng/mL) • 1999-2004: 2.9 ng/mL (WHO corrected value; the nominal value was 2.5 ng/mL) • 2005 onwards: 2.5 ng/mL • Positive PSA test led to DRE and TRUS and laterally directed sextant biopsies
Frequency	• Every 2 years until the men reached the upper age limit for invitation to screening; during screening round four, men born between 1930 and 1931 were no longer invited, and during screening round five, men born between 1932 and 1933 were not invited, and so forth

	(Bergdahl et al, 2009) - Mean age at last invitation to screening: 69 years (67-71) • Seven screening rounds were completed by the end of 2008			
Comparator				
Description of comparator	Men not invited for screening			
Outcomes				
Primary outcome	Prostate cancer-specific mortality			
Secondary outcome	Incidence of prostate cancer			
Other outcomes	N/A			
Key results				
Population analysed	**Screening**		**Control**	
Randomised	10,000		10,000	
Efficacy analysis (ITT)	9952		9952	
Efficacy analysis (PP)	NR		NR	
Safety analysis	NR		NR	
Outcome	**Control n/N (%)**	**Screening (Total) n/N (%)**	**Screening (Attendees) n/N (%)**	**Screening (Non-attendees) n/N (%)**
Incidence of prostate cancer NB. Relative risks not reported				
Diagnosis of prostate cancer	718/9952 (7.2)	1138/9952 (11.4)	1046/7578 (13.8)	92/2374 (3.9)
Low risk[a]	199/9952 (2.0)	604/9952 (6.1)	590/7578 (7.8)	14/2374 (0.6)
Moderate risk[b]	249/9952 (2.5)	363/9952 (3.6)	339/7578 (4.5)	24/2374 (1.0)
High risk[c]	126/9952 (1.3)	96/9952 (1.0)	76/7578 (1.0)	20/2374 (0.8)
Advanced disease[d]	87/9952 (0.9)	46/9952 (0.5)	25/7578 (0.3)	21/2374 (0.9)
	• Significantly lower advanced prostate cancer in the screening group compared with the control • P=0.0003			
Unknown[e]	57/9952 (0.6)	29/9952 (0.3)	16/7578 (0.2)	13/2374 (0.5)
All-cause mortality according to birth cohort at entry to study NB. Relative risks not reported				
Total	1982/9952 (19.92)	1981/9952 (19.91)	1115/7578 (14.71)	866/2374 (36.48)
1930-1934 (60-64 years)	853/2789 (30.58)	836/2774 (30.14)	488/2064 (23.64)	348/710 (49.01)
1935-1939 (55-59 years)	650/3161 (20.56)	634/3123 (20.30)	360/2420 (14.88)	274/703 (38.98)
1940-1944 (50-54 years)	479/4002 (11.97)	511/4055 (12.60)	267/3094 (8.63)	244/961 (25.39)
Prostate cancer-specific mortality				
Total	78/9952 (0.78)	44/9952 (0.44)	27/7578 (0.36)	17/2374 (0.72)
	• Relative risk of dying from prostate cancer - RR 0.56; 95% CI 0.39-0.82 - Favours screening - P=0.002 • Secondary analysis: relative risk of death from prostate cancer for attendees compared to the control group - RR 0.44; 95% CI 0.28-0.68			

	- Favours screening - P=0.0002 • Secondary analysis: relative risk of death from prostate cancer for non-attendees compared to the control group - RR 1.05; 95% CI 0.62-1.78 - No difference - P=0.84			
1930-1934 (60-64 years)	35/2789 (1.25)	27/2774 (0.97)	19/2064 (0.92)	8/710 (1.13)
1935-1939 (55-59 years)	35/3161 (1.11)	12/3123 (0.38)	6/2420 (0.25)	6/703 (0.85)
1940-1944 (50-54 years)	8/4002 (0.20)	5/4055 (0.12)	2/3094 (0.06)	3/961 (0.31)
Other results				
Prostate cancer-specific metastatic disease	See "diagnosis of prostate cancer results for advanced disease" (above)			
Skeletal-related events	NR			
Quality of life	NR			
Test performance characteristics	NR			
Harms and benefits of biopsy	NR			
Harms and benefits of treatment	NR			
Other harms and benefits	NR			
Internal Validity				
Source of quality assessment	Evidence reviewers			
Overall quality rating	Fair			
Comments	• Results from the 1930-34 and 1935-39 cohort have been reported as the Swedish arm of the overall ERSPC trial			
External validity				
Generalisability	Age range of participants: 50-64 years			
Applicability	RCT took place in Goteborg, Sweden [PSA test cut-off varied over time (1995-1998: 3.4 ng/mL (WHO corrected value; the nominal value was 3.0 ng/mL); 1999-2004: 2.9 ng/mL (WHO corrected value; the nominal value was 2.5 ng/mL); 2005 onwards: 2.5 ng/mL]			
Comments - Contamination?	- Contamination in control group: Not specified. Only reported as "low" - Compliance rate for screening: 7578/9952 (76%) - Compliance rate for prostate biopsy: 2298/2469 (93%)			
Additional references				
Bergdahl A, Aus G, Lilja H, and Hugosson J. (2009). Risk of dying from prostate cancer in men randomised to screening. *Cancer* 115(24): 5672-5679.				

Abbreviations: CI, confidence interval; DRE, digital rectal examination; N/A, not applicable; NR, not reported; PSA, prostate-specific antigen; RCT, randomised controlled trial; RR, relative risk; TRUS, transrectal ultrasound.

a T1, not N1 or M1, and Gleason score ≤6 and PSA value <10 ng/mL

b T1-2, but not N1 or M1, with a Gleason score ≤7, PSA value <20 ng/mL or both; and not meeting the criteria for low risk

c T1-4, but not N1 or M1, with a Gleason score ≥8, PSA value <100 ng/mL or both; and not meeting the criteria for low or moderate risk

d N1 or M1, or PSA value ≥100 ng/mL

e Includes seven cases detected at autopsy

Study ID	Norrkoping
Citation/Primary publication	Sandblom G, Varenhorst E, Rosell J, Lofman O and Carlsson P. (2011). Randomised prostate cancer screening trial: 20 year follow-up. *BMJ* 342: d1539.
Affiliation/Source of funds	• Research Council of the South-East Region of Sweden • Swedish Cancer Foundation • County Council of Ostergotland
Study characteristics	
Study design	• RCT in Norrkoping, Sweden that commenced in 1987 • All men aged between 50-69 years residing in Norrkoping were identified through the Swedish National Population Register • 1494 eligible men were randomly allocated to be screened and invited to participate in the prostate cancer screening trial • This study reports on up to 20-years of follow-up of participants
Randomisation method	Every 6th man was randomly allocated to the screening group from a list of date of births obtained from the Swedish National Population Register. The remaining men served as controls
Location (country)	Norrkoping, Sweden
Inclusion criteria	Men aged between 50-69 years living in Norrkoping, Sweden
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	Men with an earlier diagnosis of prostate cancer
Study size	• No. eligible men aged 50-69 years in the population: 9026 • 1494 men randomly selected and invited for screening - 1st round in 1987: 1161/1494 (78%) underwent screening - 2nd round in 1990: 957/1363 (70%) - 3rd round in 1993: 895/1210 (74%) - 4th round: 512 men excluded for being born before 1927. 446/606 (74%) remaining men underwent screening • Remaining 7532 men served as a control group
Length of follow-up *Mean/median (years)*	• Median: 6.3 • Maximum length of follow-up: 20
Population	
Age	• 50-69 years • Only men 69 years or younger were invited to the 4th screening round in 1996
Risk factors	NR
Comorbidities	NR
Other key population characteristics	• No. men in screening group: 1494 - 50-54 years: 357 - 55-59 years: 360 - 60-64 years: 393 - 65-69 years: 384 • No. men in no screening group: 7532 - 50-54 years: 1790 - 55-59 years: 1809 - 60-64 years: 2004 - 65-69 years: 1929
Intervention	
Description of test +/- DRE +/- TRUS	• 1st and 2nd rounds of screening were performed by DRE only • 3rd and 4th rounds of screening included DRE + PSA • Participants who had abnormal findings on DRE and/or PSA >4.0 ng/mL underwent TRUS guided biopsies
PSA test cut-off	4.0 ng/mL
Frequency	Every 3 years with 4 screenings in total
Comparator	
Description of comparator	Men not invited for screening

<table>
<tr><th colspan="5">Outcomes</th></tr>
<tr><td>Primary outcome</td><td colspan="4">Prostate cancer mortality at 20 years follow-up</td></tr>
<tr><td>Secondary outcome</td><td colspan="4">N/A</td></tr>
<tr><td>Other outcomes</td><td colspan="4">• Clinical stage
• Choice of therapy in men diagnosed with prostate cancer across both screened and control groups
• Number of prostate cancers diagnosed</td></tr>
<tr><th colspan="5">Key results</th></tr>
<tr><th>Population analysed</th><th colspan="2">Intervention</th><th colspan="2">Comparator</th></tr>
<tr><td>Randomised</td><td colspan="2">1494</td><td colspan="2">7532</td></tr>
<tr><td>Efficacy analysis (ITT)</td><td colspan="2">NR</td><td colspan="2">NR</td></tr>
<tr><td>Efficacy analysis (PP)</td><td colspan="2">NR</td><td colspan="2">NR</td></tr>
<tr><td>Safety analysis</td><td colspan="2">NR</td><td colspan="2">NR</td></tr>
<tr><td>Diagnosis of prostate cancer</td><td colspan="4">• 85/1494 (5.7%) men in the screening group were diagnosed with prostate cancer
- 43 (2.9%) were found at screening
- 42 (2.58%) were found in the interval between examinations
• 292/7532 (3.9%) men in the control group were diagnosed with prostate cancer
• Mean (SD) age at diagnosis (years)
- Screening group: 68.1 (5.6)
- 66.5 (5.2) in men with cancer detected at screening
- 69.9 (5.5) in men with interval cancers
- Control: 69.7 (5.7)
• Men with localised tumours (T1-2, N0/NX, M0) (%)
- Screening group: 48/85 (56.5)
- Control group: 78/292 (26.7)
- P<0.001
• Men with advanced tumours (T3-4. N1 or MX/M1) (%)
- Screening group: 37/85 (43.5)
- Control group: 213/292 (73.3)
• Men with non-localised tumours (%)
- Screening group: 37/1494 (2.5%)
- Control group: 213/7532 (2.8%)
- P=0.44</td></tr>
<tr><th>Outcome</th><th>Screening
n/N (%)</th><th>Control
n/N (%)</th><th>Relative risk estimate
(95% CI)</th><th>P-value
Favours screening/control</th></tr>
<tr><td rowspan="2">Prostate cancer-specific mortality <u>in men with prostate cancer</u></td><td>30/85
(35)</td><td>130/292
(45)</td><td>1.16 (0.78-1.73)</td><td>No difference</td></tr>
<tr><td colspan="4">• Hazard ratios:
- Death from prostate cancer: 1.23 (0.94-1.62); P=0.13
- Death from prostate cancer after adjustment for age at start of the study: 1.58 (1.06-2.36); P=0.024
- Death from prostate cancer with the addition of the period of diagnosis as a dichotomous time dependent variable and adjustment for age at start of the study: 1.59 (1.07-2.38); P=0.022</td></tr>
<tr><td>All-cause mortality <u>in men with prostate cancer</u></td><td>69/85
(81)</td><td>252/292
(86)</td><td>NR</td><td>NR</td></tr>
<tr><td>Prostate cancer-specific metastatic disease <u>in men with prostate cancer</u></td><td colspan="4">• Men with advanced tumours (T3-4. N1 or MX/M1) (%)
- Screening group: 37/85 (43.5)
- Control group: 213/292 (73.3)</td></tr>
<tr><td>Skeletal-related events</td><td colspan="4">NR</td></tr>
<tr><td>Quality of life</td><td colspan="4">NR</td></tr>
<tr><td>Other results</td><td colspan="4">• Median cancer-specific survival (months)
- Screened group: 201
- Control group: 133
• No significant difference in prostate cancer survival (P=0.065) or overall survival (P=0.14) for</td></tr>
</table>

	men with prostate cancer diagnosed in the screening group compared with the men diagnosed with prostate cancer in the control group
Other results	
Test performance characteristics	NR
Harms and benefits of biopsy	NR
Harms and benefits of treatment	NR
Other harms and benefits	NR
Internal Validity	
Source of quality assessment	Evidence reviewer
Overall quality rating	Poor
Comments	This is a pseudo-RCT and thus Level III-1 evidence
External validity	
Generalisability	Age range of participants: 50-69 years
Applicability	The RCT took place in Norrkoping, Sweden. PSA test cut-off 4.0 ng/mL
Comments - Contamination?	• Contamination in control group: not specified but the authors report on a "low rate of contamination" • Compliance rate for screening: 70-78% depending on year • Compliance rate for prostate biopsy: 304/310 (98%) (Sandblom et al, 2004)
Additional references	
Sandblom G, Varenhorst E, Lofman O, Rosell J, Carlsson P. (2004) Clinical consequences of screening for prostate cancer: 15 years follow-up of a randomised controlled trial in Sweden. *Eur Urol* 46:717-724.	

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ITT, intention-to-treat; N/A, not applicable; NR, not reported; PP, per-protocol; PSA, prostate-specific antigen; RCT, randomised controlled trial; TRUS, transrectal ultrasound.

Study ID	Stockholm
Citation/Primary publication	Kjellman A, Akre O, Norming U, Tornblom M, and Gustafsson O. (2009). 15-year follow up of a population based prostate cancer screening study. *J Urol* 181:1615-1621.
Affiliation/Source of funds	• Stockholm County Council • The Thure and Brita Grafstrom Foundation • A grant from Odd Fellows 164 Sodertalje • A grant from the Stockholm County Council and Karolinska Institutet
Study characteristics	
Study design	• RCT in Stockholm, Sweden that commenced in 1988 • All men aged between 55-70 years with current addresses in the catchment area of Stockholm South Hospital were identified through the Swedish census records • 2400 eligible men were randomly selected and invited to participate in a prostate cancer screening study • This study reports on a 15-year follow-up of participants
Randomisation method	Not specified. The authors only state that patients were randomly selected for screening. No additional information is provided on the method of randomisation
Location (country)	Stockholm, Sweden
Inclusion criteria	Men aged between 55-70 years living in the catchment area of Stockholm South Hospital
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	Men with an earlier diagnosis of prostate cancer
Study size	<u>As per "Materials and Methods"</u> • No. eligible men aged 55-70 years in the population: 26,602/27,204[a] ("source population") • 2400 men randomly selected and invited to participate in the study ("invited participants") - 1782 (74%) accepted the invitation and received screening ("attendees") - 580 did not attend, 17 were not reachable by mail and 21 had previously diagnosed prostate cancer that was not detected at the screening exclusion round. Thus, 618 (26%) men comprised the "non-attendees" • The 24,202/24,804[a] remaining men served as a control group and received usual care <u>As per "Results"</u> • Number of subjects: - Total no. men followed: 27,146 - Source population: 24,772 - Invited participants: 2374 - Attendees: 1769 - Non-attendees: 605
Length of follow-up *Mean/median (years)*	• Median (range): 12.9 (0.2-15.7) • Maximum length of follow-up: 15.7 • Mean (range): - Source population: 13.0 (0.7-15.7) - Invited participants: 12.9 (0.2-15.7) - Attendees: 13.3 (0.2-15.7) - Non-attendees: 11.8 (0.2-15.7)
Population	
Age	• 55-70 years • Mean age in years at study start (range) - Source population: 62.3 (54.3-70.2) - Invited participants: 62.4 (54.1-70.2) - Attendees: 62.5 (54.3-70.2) - Non-attendees: 62.0 (54.1-70.2)
Risk factors	NR
Comorbidities	NR
Other key population characteristics	None
Intervention	
Description of test +/- DRE +/- TRUS	• PSA + DRE + TRUS • Participants who had abnormal findings on DRE and/or TRUS underwent TRUS guided biopsies

PSA test cut-off	• PSA >7.0 ng/mL = repeat TRUS performed • PSA >10.0 ng/mL = randomised quadrant biopsies taken • With this protocol, 65 patients (3.6%) with prostate cancer were diagnosed among all screened men • DRE and/or TRUS initiated biopsies detected 62 of the 65 cases • The remaining 3 cancers were detected after repeat TRUS or after biopsies following increased PSA
Frequency	Once only (single screening)
Comparator	
Description of comparator	Eligible men not invited for screening
Outcomes	
Primary outcome	Prostate cancer mortality at 15 years of follow-up
Secondary outcome	None
Other outcomes	• Any cause mortality (including attendees and non-attendees) • Other cause mortality (including attendees and non-attendees) • Number of prostate cancers diagnosed

Key results

Population analysed	Intervention	Comparator
Randomised	2400	24,202/24,804[a]
Efficacy analysis (ITT)	2374	24,772[a]
Efficacy analysis (PP)	NR	NR
Safety analysis	NR	NR

Characteristic	Source population	Invited participants	Attendees	Non-Attendees
Prostate cancer diagnosis				
No. subjects	**24,772**	**2374**	**1769**	**605**
No. prostate cancer /No. subjects (%)	1972/24,772 (8.0)	208/2374 (8.8)	169[b]/1769 (9.6)	39/605 (0.1)
No. prostate cancer/1000 PYFU (95% CI)	5.2 (5.0-5.4)	4.0 (3.4-4.7)	3.9 (3.3-4.8)[c]	4.1 (3.0-5.7)
Mean age at prostate cancer diagnosis (range)	72.8 (56.0-84.9)	70.3 (55.0-83.0)	70.4 (55.0-83.0)	70.0 (57.0-83.0)
Prostate cancer-specific mortality				
No. deaths/No. prostate cancer (%)	506/1972 (28)	53/208 (25.5)	38/169[b] (22.5)	15/39 (38.5)
No. 1000 person years (95% CI)	1.57 (1.44-1.71)	1.72 (1.32-2.26)	1.61 (1.17-2.21)	2.11 (1.27-3.50)
IRR (95% CI)[d]	1 (referent)	1.10 (0.83-1.46)	1.04 (0.76-1.45)	1.28 (0.76-2.13)
Other cause death/all-cause mortality				
No. any cause death/No. subjects (%)	10,328/24,772 (41.7)	986/2374 (41.5)	648/1769 (36.6)	338/605 (55.9)
No. other cause deaths (excluding prostate cancer death)/No. subjects (%)	9822/24,772 (39.7)	933/2374 (39.3)	610/1769 (34.5)	323/605 (53.4)
No. 1000 person years (95% CI)	30.5 (29.9-31.1)	30.4 (28.5-32.4)	25.8 (23.8-28.0)	45.4 (40.7-50.7)
IRR (95% CI)[d]	1 (referent)	0.98 (0.92-1.05)	0.82 (0.76-0.90)	1.53 (1.37-1.71)
IRR (95% CI)	NR	NR	1 (referent)	1.89 (1.65-2.16)

Relative risk of death from causes other than prostate cancer stratified by follow-up

	Overall IRR		Follow-up IRR (95% CI)			
	Crude	Age-adjusted	0-1 year	1-5 years	5-10 years	10+ years
Source population	1	1	1	1	1	1
Invited participants	1.01	1.02 (0.95-1.09)	0.68 (0.41-1.13)	0.94 (0.82-1.08)	1.04 (0.92-1.18)	1.05 (0.95-1.16)
Attendees	0.87	0.82 (0.76-0.89)	0.46 (0.17-1.25)	0.90 (0.76-1.07)	0.81 (0.70-0.94)	0.90 (0.76-1.07)
Non-attendees	1.35	1.55 (1.39-	4.58 (2.58-	1.56 (1.24-	1.46 (1.20-	1.53 (1.29-

		1.74)	8.11)	1.96)	1.79)	1.81)
No. deaths		10,755	137	2341	3352	4925
Other results	• There was no significant difference in survival from prostate cancer between source and invited populations (log rank p=0.87) • There was a significant difference in prostate cancer-specific survival between attendees and source population (log rank p=0.000). The screening attendees had significantly improved survival • There was a significant difference in survival until death from causes other than prostate cancer between attendees and non-attendees (log rank p=0.000)					
Other results						
Prostate cancer-specific metastatic disease	NR					
Skeletal-related events	NR					
Quality of life	NR					
Test performance characteristics	NR					
Harms and benefits of biopsy	NR					
Harms and benefits of treatment	NR					
Other harms and benefits	NR					
Internal Validity						
Source of quality assessment	Evidence reviewer					
Overall quality rating	Poor					
Comments	None					
External validity						
Generalisability	Age range or participants: 55-70 years					
Applicability	RCT took place in Stockholm, Sweden. Screening intervention involved PSA + DRE + TRUS. Participants who had abnormal findings on DRE and/or TRUS underwent TRUS guided biopsy. PSA >7.0 ng/mL led to repeated TRUS performed. PSA >10.0 ng/mL resulted in biopsy					
Comments - Contamination?	• Contamination in control group: NR • Compliance rate for screening: 74% • Compliance rate for prostate biopsy: NR • There was a discrepancy between population sizes because the file containing the registration numbers of the original cohort could not be retrieved. When the cohort was reconstructed, an additional 602 (2%) registration numbers were found compared to the original source population					

Abbreviations: CI, confidence interval; DRE, digital rectal examination; IRR, incidence rate ratio, ITT, intention-to-treat; NR, not reported; PP, per-protocol; PSA, prostate-specific antigen; PYFU, person years at follow-up; RCT, randomised controlled trial; TRUS, transrectal ultrasound.

a Report has internal discrepancies about this number because the file containing the registration numbers of the original cohort could not be retrieved.

b Includes the 65 screening detected cases.

c Excludes the 65 prevalent cases detected by screening.

d Adjusted for attained age as a time dependent variable.

Study ID	Quebec
Citation/Primary publication	Labrie F, Candas B, Cusan L, Gomez JL, Belanger A, Brousseau g, Chebrette E and Levesque J. (2004). Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. *Prostate* 59(3): 311-318.
Affiliation/Source of funds	• Laval University Medical Centre • Laval University, Quebec, Canada
Study characteristics	
Study design	• RCT in Quebec, Canada that commenced in 1988 • All men aged between 45-80 years who were registered in the electoral roll of the Quebec city area were identified • Eligible men were randomised to either the group invited for prostate cancer screening or to the control group not invited for screening at a ratio of 2:1 in favour of screening • This study reports on a 11-year follow-up of participants
Randomisation method	Not specified. The authors only state that the men were randomly allocated either to the group invited for annual screening or to the control group not invited for screening at a ratio of 2:1 in favour of screening. No additional information is provided on the method of randomisation
Location (country)	Quebec, Canada
Inclusion criteria	Men aged between 45-80 years registered in the electoral roll of the Quebec city area
Definition of *asymptomatic*	Not specifically defined
Exclusion criteria	• Men with a diagnosis of prostate cancer before 15 November 1988 • Men who had previous screening and were referred to the study clinic for consultation
Study size	• No. eligible men aged 45-80 years in the population: 46,486 • 31,133 men invited to participate in the study after randomisation - 7348 (23.6%) underwent screening - 23,785 unscreened men • Remaining 15,353 men served as a control group - 1122 (7.3%) received screening - 14,231 unscreened men
Length of follow-up *Mean/median (years)*	• Median follow-up duration of screened men: 7.93 • Maximum length of follow-up: 11
Population	
Age	• 45-80 years • Median age in years (Q1-Q3) - Invited, screened men: 60 (55.2-65.8) - Invited, unscreened men: 59.7 (52.5-67.8) - Not invited, screened men: 61.1 (56.5-66.2) - Not invited, unscreened men: 59.0 (52.0-66.8)
Risk factors	NR
Comorbidities	NR
Other key population characteristics	N/A
Intervention	
Description of test +/- DRE +/- TRUS	• First screening round: PSA + DRE • Follow-up screenings: PSA only
PSA test cut-off	• First screening round: PSA + DRE - TRUS performed if PSA >3.0 ng/mL and/or abnormal DRE except for the first 1002 who had PSA + DRE + TRUS performed • Follow-up screenings: PSA only - TRUS performed if PSA >3.0 ng/mL for the first time - If PSA already >3.0 ng/mL at a previous visit, TRUS was only performed if PSA had increased by more than 20% compared with the value measured 1 year earlier or if the serum PSA had increased more than 20% over the predicted PSA if calculated at a previous visit

Frequency	• Annually • Median delay to first screening in the invited, screened group: 3.19 years			
Comparator				
Description of comparator	Eligible men not invited for screening			
Outcomes				
Primary outcome	Prostate cancer mortality at 11 years of follow-up			
Secondary outcome	N/A			
Other outcomes	• Prostate cancer death incidence rates in screened versus unscreened cohorts • Clinical stage and choice of therapy in men diagnosed with prostate cancer			
Key results				
Population analysed	**Intervention**		**Comparator**	
Randomised	31,133		15,353	
Efficacy analysis (ITT)	NR		NR	
Efficacy analysis (PP)	NR		NR	
Safety analysis	NR		NR	
Outcome	**Screening n/N (%)**	**Control n/N (%)**	**Risk estimate (95% CI)**	**P-value** *Favours screening/control*
Prostate cancer mortality				
Total	10/7348 (0.14)	74/14,231 (0.52)	1.01 (0.82-1.40)	No difference
Adjusted for difference in length of follow up[a]	10/7348 (0.14)	74/14,231 (0.52)	0.49 (0.25-0.99)	Favours screening P=0.047
Adjusted for contamination[b]	11/8470 (0.13)	217/38,016 (0.57)	0.36 (0.19-0.65)	Favours screening P<0.0002
Analysis of the effect screening, group at randomisation and age on prostate cancer mortality using the Cox Proportional Hazards Models	• Screened versus unscreened - RR: 0.385 (0.207-0.714); P=0.0025 • Invited versus not invited - RR: 1.085 (0.822-1.433); P=0.5637 • Age (on 15 November 1988) - RR: 0.938 (0.897-0.981); P=0.0054			
Related results	• Exposures - Invited screened group: 50,433 man-years - Control unscreened group: 141,535 man-years • Annual cause-specific death rate incidences over the 11-year period: - Invited screened group: 19.8 per 100,000 man-years - Control unscreened group: 52.3 per 100,000 man-years - P <0.002 (prostate cancer death rate incidence is 62% lower in the screened men versus the control group)			
Other results				
All-cause mortality	NR			
Prostate cancer-specific metastatic disease	1/159 cancers (0.6%) diagnosed at follow-up visits was metastatic, thus permitting 99.4% of patients to be diagnosed at a localised stage			
Skeletal-related events	NR			
Quality of life	NR			
Test performance characteristics	NR			
Harms and benefits of biopsy	NR			
Harms and benefits of treatment	NR			
Other harms and benefits	NR			
Internal Validity				
Source of quality assessment	Evidence reviewer			
Overall quality rating	Poor			
Comments	None			

External validity	
Generalisability	Age range of participants: 45-80 years
Applicability	The RCT took place in Quebec, Canada. PSA test cut-off 3.0 ng/mL
Comments - Contamination?	• Contamination in control group: 1122/15,353 (7.3%). However, this only reflects men in the control group who came on their own to the study site clinic for screening. There was no report of any other withdrawals or whether participants in the control group were screened somewhere other than the study site. Hence, it is possible that more than 7.3% of the control group were actually screened • Compliance rate for screening: 23.6% • Compliance rate for prostate biopsy: NR

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ITT, intention-to-treat; N/A, not applicable; NR, not reported; PP, per-protocol; PSA, prostate-specific antigen; RCT, randomised controlled trial; RR, relative risk; TRUS, transrectal ultrasound.

a To assess a possible effect of the difference of length of follow-up between screened and unscreened men, the duration of exposure of unscreened men was adjusted to limit their follow-up to a maximum of 7.93 years and therefore match the median follow-up duration of screened men.

b Assessed by analysing the data with respect to the intervention (screened and not screened) rather than the original group of randomisation.

1.9 Evidence Statement Forms

Key question(s): Does PSA testing, with or without digital rectal examination, in asymptomatic men reduce <u>prostate cancer-specific mortality</u>?		**Evidence table ref:** Evidence Evaluation Report Tables 6 and 7
1. Evidence base *(number of studies, level of evidence and risk of bias in the included studies)*		
• Three Level I studies - **Key review**: **Ilic (2013):** good quality [included the PLCO, ERSPC, Norrkoping, Stockholm and Quebec trials] - Djulbegovic (2010): good quality [PLCO, ERSPC, Goteborg, Norrkoping, Quebec] - Lumen (2012): good quality [PLCO, ERSPC, Goteborg, Norrkoping, Stockholm, Quebec, Rotterdam-Ireland] • Five Level II studies[a] - PLCO: good quality; 13 years of follow-up - ERSPC: fair quality; 11 years of follow-up - Goteborg: fair quality; 14 years of follow-up - Stockholm: poor quality; 15 years of follow-up - Quebec: poor quality; 11 years of follow-up • One Level III-1 study[a] - Norrkoping: poor quality; up to 20 years of follow-up	A	**One or more level I studies with a low risk of bias or several level II studies with a low risk of bias**
	B	**One or two Level II studies with a low risk of bias or SR/several Level III studies with a low risk of bias**
	C	**One or two Level III studies with a low risk of bias or Level I or II studies with a moderate risk of bias**
	D	**Level IV studies or Level I to III studies/SRs with a high risk of bias**
2. Consistency *(if only one study was available, rank this component as 'not applicable')*		
• All three systematic reviews concluded that there was no significant effect of prostate cancer screening on prostate cancer-specific mortality - Ilic (2013): - Including all enrolled men in the ERSPC trial: [RR 1.00; 95% CI 0.86-1.17]; moderate heterogeneity; 341,342 total patients - All enrolled men excluding Norrkoping, Stockholm & Quebec: [RR 0.96; 95%CI 0.70-1.30] ; substantial heterogeneity; 258,684 total patients - Including the core age group of men (55-69 years) in the ERSPC trial: [RR 1.00; 95% CI 0.83-1.19]; substantial heterogeneity; 321,586 total patients - Core age group of men excluding Norrkoping, Stockholm & Quebec: [RR 0.94; 95%CI 0.65-1.35] ; substantial heterogeneity; 238,928 total patients - Lumen (2012): [RR 0.88; 95% CI 0.72-1.06]; substantial heterogeneity; 486,813 total patients - Djulbegovic (2010): [RR 0.88; 95% CI 0.71-1.09]; substantial heterogeneity;	A	**All studies consistent**
	B	**Most studies consistent and inconsistency can be explained**
	C	**Some inconsistency, reflecting genuine uncertainty around question**
	D	**Evidence is inconsistent**
	NA	**Not applicable (one study only)**

302,500 total patients • Level II studies - PLCO: [RR 1.09; 95% CI 0.87-1.36] - ERSPC (all enrolled men): [RR 0.83; 95% CI 0.72-0.94]; favours screening - ERSPC (core age group of men): [RR 0.79; 95% CI 0.68-0.91]; favours screening - Goteborg: [RR 0.56; 95% CI 0.39-0.82]; favours screening - Stockholm: [RR 1.10; 95% CI 0.83-1.46] - Quebec: [RR 1.01; 95% CI 0.82-1.40] • Level III-1 study - Norrkoping: [RR 1.16; 95% CI 0.78-1.73]		
3. Clinical impact[b] *(Indicate if the study results varied according to some <u>unknown</u> factor (not simply study quality or sample size) and thus the clinical impact of the intervention could not be determined)*		
• Prostate cancer-specific mortality (%) in the screening and control population - Ilic (2013): - Including all enrolled men in the ERSPC trial: screening 0.45%, control 0.71% - Including the core age group of men in the ERSPC trial: screening 0.43%, control 0.72% - Lumen (2012): screening 0.34%, control 0.56% - Djulbegovic (2010): not reported	A	**Very large**
	B	**Substantial**
	C	**Moderate**
	D	**Slight/Restricted**
	NA	**Not applicable/no difference/underpowered**
4. Generalisability[b] *(How well does the body of evidence match the population and clinical settings being targeted by the Guideline?)*		
• Age range of participants in each RCT - PLCO: 55-74 years - ERSPC: variable between 50-74 years according to study site but all sites included a predefined core age group of men aged 55-69 years - Goteborg: 50-64 years - Norrkoping: 50-69 years - Stockholm: 55-70 years - Quebec: 45-80 years	A	**Evidence directly generalisable to target population**
	B	**Evidence directly generalisable to target population with some caveats**
	C	**Evidence not directly generalisable to the target population but could be sensibly applied**
	D	**Evidence not directly generalisable to target population and hard to judge whether it is sensible to apply**
5. Applicability[b] *(Is the body of evidence relevant to the Australian healthcare context in terms of health services/delivery of care and cultural factors?)*		
• Country where each RCT took place - PLCO: United States of America - ERSPC: eight European countries - Goteborg: Sweden - Norrkoping: Sweden - Stockholm: Sweden - Quebec: Canada	A	**Evidence directly applicable to Australian healthcare context**
	B	**Evidence applicable to Australian healthcare context with few caveats**
	C	**Evidence probably applicable to Australian healthcare context with some caveats**
	D	**Evidence not applicable to Australian healthcare context**

Other factors *(Indicate here any other factors that you took into account when assessing the evidence base (for example, issues that might cause the group to downgrade or upgrade the recommendation)*

- The significant effect of screening on prostate cancer-specific mortality in the ERSPC trial was affected more by Sweden than any other country. This Swedish arm comprises all of the participants in the Goteborg trial which featured younger participants, a lower PSA cut-off, shorter screening intervals and a longer follow-up.
- Differences in intervention (PSA test +/- DRE +/- TRUS) between the RCTs.
- Differences in PSA test cut-off between the RCTs.
- High contamination rate in the PLCO and ERSPC trials, which the EAG noted can result in lower than expected mortality.
- The EAG commented that the mortality rates of the individual countries within ERSPC were very different and there was poor compliance for biopsy. Furthermore, single rather than multiple biopsy was undertaken in the RCTs, which makes them less applicable to current practice.
- The Rotterdam-Ireland trial (in Lumen et al, 2012) is not a prospective RCT. Rather, it was a comparison between a screened population (part of the Rotterdam section of the ERSPC trial) and a population where screening is not routinely carried out (Ireland).
- When rating the evidence base, the EAG decided to focus on the quality of the Level II studies, rather than the quality of the Level I studies. Although the quality of each of the Level I studies was good, the quality of the individual studies within them was considered to be variable.
- The EAG noted that all of the systematic reviews showed substantial heterogeneity in their meta-analyses of prostate cancer-specific mortality .The EAG could not rule out a small effect in either direction (i.e. a small increase or a small decrease).

EVIDENCE STATEMENT MATRIX

Please summarise the development group's synthesis of the evidence relating to the key question, taking all the above factors into account.

Component	Rating	Description
1. Evidence base	B	Three Level I studies, comprising a total of five Level II studies (one of good quality, two of fair quality and two of poor quality)[a] and one Level III-1 study of poor quality.
2. Consistency	C	All of the systematic reviews showed substantial heterogeneity in their meta-analysis of prostate cancer-specific mortality. There is thus some inconsistency, which is likely due to methodological and quality differences between the included studies.

EVIDENCE STATEMENT

In asymptomatic men, the present evidence is inconsistent as to whether there is an effect of PSA testing, with or without DRE, on the risk of prostate cancer-specific mortality compared with no PSA testing, although the possibilities of no effect or a small protective effect cannot be excluded.

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; PLCO, Prostate, Lung and Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; RR, relative risk; SR, systematic review; TRUS, transrectal ultrasound.

[a] The quality rating of the Level II and Level III-1 studies should be considered together with the limitations of each study as reported in Section 2.3 of the Evidence Evaluation Report.

[b] At a meeting held on 22 January 2013, it was agreed by the NHMRC and EAG that the clinical impact, generalisability and applicability components would not require rating for the purposes of this evaluation.

Key question(s): Does PSA testing, with or without digital rectal examination, reduce <u>all-cause mortality</u> in asymptomatic men?		**Evidence table ref:** Evidence Evaluation Report Tables 9 and 10
1. Evidence base *(number of studies, level of evidence and risk of bias in the included studies)*		
• Three Level I studies with meta-analysis - **Key review**: **Ilic (2013):** good quality [included the PLCO, ERSPC, Norrkoping and Stockholm trials] - Djulbegovic (2010): good quality [PLCO, ERSPC, Goteborg, Norrkoping] - Lumen (2012): good quality [PLCO, Goteborg, Stockholm, Rotterdam-Ireland] • Four Level II studies[a] - PLCO: good quality; 13 years of follow-up - ERSPC: fair quality; 11 years of follow-up - Goteborg: fair quality; 14 years of follow-up - Stockholm: poor quality; 15 years of follow-up	A	**One or more Level I studies with a low risk of bias or several Level II studies with a low risk of bias**
	B	**One or two Level II studies with a low risk of bias or SR/several Level III studies with a low risk of bias**
	C	**One or two Level III studies with a low risk of bias or Level I or II studies with a moderate risk of bias**
	D	**Level IV studies or Level I to III studies/SRs with a high risk of bias**
2. Consistency *(if only one study was available, rank this component as 'not applicable')*		
• All three systematic reviews concluded that there was no significant effect of prostate cancer screening on all-cause mortality - Ilic (2013): - Including all enrolled men in the ERSPC trial: [RR 1.00; 95% CI 0.96-1.03]; substantial heterogeneity; 294,856 total patients - All enrolled men excluding Norrkoping & Stockholm: [RR 0.99; 95% CI 0.96-1.02]; no heterogeneity; 258,684 total patients - Including the core age group of men (55-69 years) in the ERSPC trial: [RR 0.99; 95% CI 0.96-1.03]; substantial heterogeneity; 275,100 total patients - Core age group of men excluding Norrkoping & Stockholm: [RR 0.98; 95% CI 0.97-1.00]; no heterogeneity; 238,928 total patients - Lumen (2012): [RR 0.90; 95% CI 0.75-1.08]; substantial heterogeneity; 269,058 total patients - Djulbegovic (2010): [RR 0.99; 95% CI 0.97-1.01]; no significant heterogeneity; 256,019 total patients • Level II studies - PLCO: [RR 0.96; 95% CI 0.93-1.00]; "borderline statistical significance in favour of screening" - ERSPC (all enrolled men): [RR 1.00; 95% CI 0.98-1.02] - ERSPC (core age group of men): [RR 0.99; 95% CI 0.97-1.01] - Goteborg: [RR 1.00; 95% CI 0.95-1.06] - Stockholm: [RR 0.98; 95% CI 0.92-1.05]	A	**All studies consistent**
	B	**Most studies consistent and inconsistency can be explained**
	C	**Some inconsistency, reflecting genuine uncertainty around question**
	D	**Evidence is inconsistent**
	NA	**Not applicable (one study only)**

3. Clinical impact[b] *(Indicate if the study results varied according to some <u>unknown</u> factor (not simply study quality or sample size) and thus the clinical impact of the intervention could not be determined)*		
• Underpowered • All-cause mortality (%) in the screening and control population - Ilic (2013): - Including all enrolled men in the ERSPC trial: screening 18.26%, control 21.07% - Including the core age group of men in the ERSPC trial: screening 17.39%, control 20.64% - Lumen (2012): screening 13.7%, control 21.1% - Djulbegovic (2010): not reported	A	**Very large**
	B	**Substantial**
	C	**Moderate**
	D	**Slight/Restricted**
	NA	**Not applicable/no difference/underpowered**
4. Generalisability[b] *(How well does the body of evidence match the population and clinical settings being targeted by the Guideline?)*		
• Age range of participants in each RCT - PLCO: 55-74 years - ERSPC: variable between 50-74 years according to study site but all sites included a predefined core age group of men aged 55-69 years - Goteborg: 50-64 years - Stockholm: 55-70 years	A	**Evidence directly generalisable to target population**
	B	**Evidence directly generalisable to target population with some caveats**
	C	**Evidence not directly generalisable to the target population but could be sensibly applied**
	D	**Evidence not directly generalisable to target population and hard to judge whether it is sensible to apply**
5. Applicability[b] *(Is the body of evidence relevant to the Australian healthcare context in terms of health services/delivery of care and cultural factors?)*		
• Country where each RCT took place - PLCO: USA - ERSPC: eight European countries - Goteborg: Sweden - Stockholm: Sweden	A	**Evidence directly applicable to Australian healthcare context**
	B	**Evidence applicable to Australian healthcare context with few caveats**
	C	**Evidence probably applicable to Australian healthcare context with some caveats**
	D	**Evidence not applicable to Australian healthcare context**

Other factors *(Indicate here any other factors that you took into account when assessing the evidence base (for example, issues that might cause the group to downgrade or upgrade the recommendation)*

- Differences in intervention (PSA test +/- DRE +/- TRUS) between the RCTs.
- Differences in PSA test cut-off between the RCTs.
- High contamination rate in the PLCO and ERSPC trials.
- The Rotterdam-Ireland trial (in Lumen et al, 2012) is not a prospective RCT. Rather, it was a comparison between a screened population (part of the Rotterdam section of the ERSPC trial) and a population where screening is not routinely carried out (Ireland).
- The Level III-1 study, Norrkoping, only reported on all-cause mortality in men diagnosed with prostate cancer and not the whole population. It is therefore not included in the Level II evidence base. Ilic (2013) included the Norrkoping study using the randomised numbers as the denominator which is inappropriate. However, they also included an analysis with the Norrkoping and Stockholm studies excluded.
- When rating the evidence base, the EAG decided to focus on the quality of the Level II studies, rather than the quality of the Level I studies. Although the quality of each of the Level I studies was good, the quality of the individual studies within them was considered to be variable.
- None of the RCTs were sufficiently powered to investigate the effect of PSA testing on all-cause mortality. The EAG agreed to carefully word the evidence statement because there could be a small effect that is not evident due to under-powering.

EVIDENCE STATEMENT MATRIX

Please summarise the development group's synthesis of the evidence relating to the key question, taking all the above factors into account.

Component	Rating	Description
1. Evidence base	B	Three Level I studies, comprising a total of four Level II studies (one of good quality, two of fair quality and one of poor quality)[a].
2. Consistency	B	The findings of all of the systematic reviews were consistent in showing no effect of prostate cancer screening on all-cause mortality. One Level II study of good quality showed borderline statistical significance in favour of screening; all other Level II studies showed no effect.

EVIDENCE STATEMENT

In asymptomatic men, PSA testing with or without DRE has no discernible effect on all-cause mortality compared with no PSA testing.

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; PLCO, Prostate, Lung and Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; RR, relative risk; SR, systematic review; TRUS, transrectal ultrasound.

[a] The quality rating of the Level II studies should be considered together with the limitations of each RCT as reported in Section 2.3 of the Evidence Evaluation Report.

[b] At a meeting held on 22 January 2013, it was agreed by the NHMRC and EAG that the clinical impact, generalisability and applicability components would not require rating for the purposes of this evaluation.

Key question(s): Does PSA testing, with or without digital rectal examination, in asymptomatic men reduce <u>prostate cancer-specific metastatic disease due to advanced prostate cancer</u>?		**Evidence table ref:** Evidence Evaluation Report Table 12 and 13
1. Evidence base *(number of studies, level of evidence and risk of bias in the included studies)*		
• Three Level I studies - **Key review**: **Ilic (2013):** good quality [included the PLCO, ERSPC, French ERSPC and Norrkoping trials] - Djulbegovic (2010): good quality [PLCO, ERSPC, French ERSPC, Norrkoping] - Lumen (2012): good quality [PLCO, ERSPC, French ERSPC, Goteborg, Norrkoping, Rotterdam-Ireland] • Three Level II studies[a] - PLCO: good quality; 13 years of follow-up - ERSPC: fair quality; 11 years of follow-up - Goteborg: fair quality; 14 years of follow-up • One Level III-1 study[a] - Norrkoping: poor quality; 20 years of follow-up	A	**One or more level I studies with a low risk of bias or several level II studies with a low risk of bias**
	B	**One or two Level II studies with a low risk of bias or SR/several Level III studies with a low risk of bias**
	C	**One or two Level III studies with a low risk of bias or Level I or II studies with a moderate risk of bias**
	D	**Level IV studies or Level I to III studies/SRs with a high risk of bias**
2. Consistency *(if only one study was available, rank this component as 'not applicable')*		
• The three systematic reviews had differing conclusions as to whether or not prostate cancer screening had an effect on prostate cancer-specific metastatic disease - Ilic (2013): - Excluding French ERSPC: [RR 0.80; 95% CI 0.73-0.87]; favours screening; no significant heterogeneity; 247,954 total patients - Including French ERSPC: [RR 0.77; 95% CI 0.71-0.83]; favours screening; heterogeneity not reported; total patients not reported - Lumen (2012): [RR 0.63; 95% CI 0.38-1.05]; substantial heterogeneity; 497,945 total patients - Djulbegovic (2010): [RR 0.94 (0.85-1.04)]; no significant heterogeneity; 332,743 total patients • Level II studies - PLCO: [Stage III prostate cancer RR 0.80; 95% CI 0.56-1.14]; [Stage IV prostate cancer RR 0.78; 95% CI 0.59-1.02] - ERSPC: [Hazard ratio 0.70; 95% CI 0.60-0.82] - Goteborg: [RR 0.53; 95% CI 0.37-0.76] • Level III-1 study - Norrkoping: [RR 0.88; 95% CI 0.62-1.24]	A	**All studies consistent**
	B	**Most studies consistent and inconsistency can be explained**
	C	**Some inconsistency, reflecting genuine uncertainty around question**
	D	**Evidence is inconsistent**
	NA	**Not applicable (one study only)**

3. Clinical impact[b] *(Indicate if the study results varied according to some unknown factor (not simply study quality or sample size) and thus the clinical impact of the intervention could not be determined)*		

• Prostate cancer-specific metastatic disease (%) in the screening and control population
- Ilic (2013):
  - Excluding French ERSPC: screening 0.77%, control 1.08%
  - Including French ERSPC: not reported
- Lumen (2012): screening 0.16%, control 0.42%
- Djulbegovic (2010): screening 0.45%, control 0.55%
	A	**Very large**

	B	**Substantial**
	C	**Moderate**
	D	**Slight/Restricted**
	NA	**Not applicable/no difference/underpowered**
4. Generalisability[b] *(How well does the body of evidence match the population and clinical settings being targeted by the Guideline?)*		
• Age range of participants in each RCT - PLCO: 55-74 years - ERSPC: variable between 50-74 years according to study site but all sites included a predefined core age group of men aged 55-69 years - Goteborg: 50-64 years - Norrkoping: 50-69 years	A	**Evidence directly generalisable to target population**
	B	**Evidence directly generalisable to target population with some caveats**
	C	**Evidence not directly generalisable to the target population but could be sensibly applied**
	D	**Evidence not directly generalisable to target population and hard to judge whether it is sensible to apply**
5. Applicability[b] *(Is the body of evidence relevant to the Australian healthcare context in terms of health services/delivery of care and cultural factors?)*		
• Country where each RCT took place - PLCO: United States of America - ERSPC: eight European countries - Goteborg: Sweden - Norrkoping: Sweden	A	**Evidence directly applicable to Australian healthcare context**
	B	**Evidence applicable to Australian healthcare context with few caveats**
	C	**Evidence probably applicable to Australian healthcare context with some caveats**
	D	**Evidence not applicable to Australian healthcare context**

Other factors *(Indicate here any other factors that you took into account when assessing the evidence base (for example, issues that might cause the group to downgrade or upgrade the recommendation)*

- Differences in the definition of what constitutes prostate cancer-specific metastatic disease. Ilic (2013) included both Stage III and IV prostate cancers from the PLCO trial in their meta-analysis. However, only Stage IV prostate cancers were defined as metastatic in the PLCO trial.
- Data specific for metastatic disease in the core age group of men in the ERSPC trial were published separately to the 2012 ERSPC trial report. This separate data was not used in the meta-analysis of advanced prostate cancer by Ilic (2013).
- Differences in intervention (PSA test +/- DRE +/- TRUS) between the RCTs.
- Differences in PSA test cut-off between the RCTs.
- High contamination rate in the PLCO and ERSPC trials.
- When rating the evidence base, the EAG decided to focus on the quality of the Level II studies, rather than the quality of the Level I studies. Although the quality of each of the Level I studies was good, the quality of the individual studies within them was considered to be variable.

EVIDENCE STATEMENT MATRIX

Please summarise the development group's synthesis of the evidence relating to the key question, taking all the above factors into account.

Component	Rating	Description
1. Evidence base	B	Three Level I studies, comprising a total of three Level II studies (one of good quality and two of fair quality) and one Level III-1 study of poor quality.[a]
2. Consistency	B	Most studies consistent and inconsistency can be explained.

EVIDENCE STATEMENT

In asymptomatic men, PSA testing with or without DRE reduces the risk of prostate cancer metastases at diagnosis compared with no PSA testing.

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; PLCO, Prostate, Lung and Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; RR, relative risk; SR, systematic review; TRUS, transrectal ultrasound.

[a] The quality rating of the Level II and Level III-1 studies should be considered together with the limitations of each study as reported in Section 2.3 of the Evidence Evaluation Report.

[b] At a meeting held on 22 January 2013, it was agreed by the NHMRC and EAG that the clinical impact, generalisability and applicability components would not require rating for the purposes of this evaluation.

Key question(s): Does PSA testing, with or without digital rectal examination, in asymptomatic men affect quality of life due to advanced prostate cancer?		**Evidence table ref:** Evidence Evaluation Report Section 2.5.3
1. Evidence base *(number of studies, level of evidence and risk of bias in the included studies)*		
No Level I or II studies of PSA-based screening were identified that reported data on quality of life in asymptomatic men with advanced prostate cancer.	A	**One or more level I studies with a low risk of bias or several level II studies with a low risk of bias**
	B	**One or two Level II studies with a low risk of bias or SR/several Level III studies with a low risk of bias**
	C	**One or two Level III studies with a low risk of bias or Level I or II studies with a moderate risk of bias**
	D	**Level IV studies or Level I to III studies/SRs with a high risk of bias**
	NA	**Not applicable**
2. Consistency *(if only one study was available, rank this component as 'not applicable')*		
Not applicable	A	**All studies consistent**
	B	**Most studies consistent and inconsistency can be explained**
	C	**Some inconsistency, reflecting genuine uncertainty around question**
	D	**Evidence is inconsistent**
	NA	**Not applicable**
3. Clinical impact[a] *(Indicate if the study results varied according to some unknown factor (not simply study quality or sample size) and thus the clinical impact of the intervention could not be determined)*		
	A	**Very large**
	B	**Substantial**
	C	**Moderate**
	D	**Slight/Restricted**
	NA	**Not applicable/no difference/underpowered**
4. Generalisability[a] *(How well does the body of evidence match the population and clinical settings being targeted by the Guideline?)*		
	A	**Evidence directly generalisable to target population**
	B	**Evidence directly generalisable to target population with some caveats**
	C	**Evidence not directly generalisable to the target population but could be sensibly applied**
	D	**Evidence not directly generalisable to target population and hard to judge whether it is sensible to apply**
5. Applicability[a] *(Is the body of evidence relevant to the Australian healthcare context in terms of health services/delivery of care and cultural factors?)*		
	A	**Evidence directly applicable to Australian healthcare context**
	B	**Evidence applicable to Australian healthcare context with few caveats**
	C	**Evidence probably applicable to Australian healthcare context with some caveats**
	D	**Evidence not applicable to Australian healthcare context**

Other factors *(Indicate here any other factors that you took into account when assessing the evidence base (for example, issues that might cause the group to downgrade or upgrade the recommendation)*

Awaiting quality of life data from key RCTs.

EVIDENCE STATEMENT MATRIX

Please summarise the development group's synthesis of the evidence relating to the key question, taking all the above factors into account.

Component	Rating	Description
1. Evidence base	NA	No Level I or Level II studies of PSA-based screening were identified that reported data on quality of life in asymptomatic men with advanced prostate cancer.
2. Consistency	NA	Not applicable.

EVIDENCE STATEMENT

It is unknown if PSA testing, with or without DRE, in asymptomatic men affects quality of life due to advanced prostate cancer, compared with no PSA testing.

Abbreviations: DRE, digital rectal examination; ERSPC, European Randomised Study of Screening for Prostate Cancer; PLCO, Prostate, Lung and Colorectal and Ovarian cancer screening trial; PSA, prostate-specific antigen; RCT, randomised controlled trial; SR, systematic review.

[a] At a meeting held on 22 January 2013, it was agreed by the NHMRC and EAG that the clinical impact, generalisability and applicability components would not require rating for the purposes of this evaluation.

1.10 Description of how comments from the EAG, NHMRC's relevant Principal Committees, Council of NHMRC and independent review have been addressed

The draft Evidence Evaluation Report and Technical Report (the draft Reports) were scrutinised by an independent reviewer to assess the methodology and ensure the review activities articulated in the draft Reports were undertaken in a transparent, accurate and unbiased manner. The draft Reports were then discussed at a meeting of the EAG held on 22 January 2013, with the content further refined and the evidence statements finalised using a consensus approach.

The revised draft Reports were circulated to EAG members on 6 February 2013 for further feedback. After minor modifications, these versions were provided for comment to NHMRC's Health Care Committee (HCC) and Prevention and Community Health Committee (PCHC) on 26 February 2013 and 28 February 2013 respectively. The resultant draft Reports were discussed again by the EAG at a meeting held on 9 April 2013, and their approval of these draft Reports achieved on 29 April 2013.

The draft Reports were again provided to PCHC and HCC on the 9 May 2013 and 14 May 2013 respectively, who were satisfied with the content and the process undertaken to develop these.

The final Evidence Evaluation Report and the Technical Report were considered by the Council of NHMRC on 21 June 2013 for recommendation to the CEO for issuing. The CEO was pleased to accept the Council's advice and agreed to issue the reports under Section 7(1a) of the National Health and Medical Research Council Act 1992.

2 Methodology for supplementary non-systematic review

2.1 Research question development

For this evaluation, the NHMRC defined a non-systematic literature review as 'a review of the published literature that does not use systematic and explicit methods to identify and include studies. They may be used to "get a flavour" for an issue but they are vulnerable to bias and should not be used as the basis for formulating recommendations. This part of the review should still include a critical appraisal and evaluation of the methodological rigour of the evidence starting with high level evidence and appraising lower levels of evidence if necessary'.

The non-systematic literature review component of this project is intended to supplement the information from the systematic review with additional evidence describing the potential harms and other benefits of PSA testing. The research topics to be addressed by the non-systematic literature review were discussed by the EAG at a meeting held on 24 August 2012, and then further developed by the evidence reviewer in consultation with the NHMRC.

The non-systematic literature review is focused on the following secondary research topics:

- Potential harms and benefits associated with PSA test performance characteristics (false positives, false negatives and the risk of overdiagnosis), including the psychological effects of PSA testing
- Potential harms and benefits associated with follow-up procedures (such as biopsy)
- Potential harms and benefits associated with active treatment options (including radical prostatectomy, radiotherapy, androgen deprivation (hormone) therapy, cryotherapy, and high intensity focused ultrasound)
- Other benefits of PSA testing

2.2 Literature search

The non-systematic review of potential harms and other benefits of PSA testing supplements the systematic review with additional evidence sourced from:

1. The Level I and Level II studies identified in the systematic review
2. Non-systematic literature searches designed to identify studies relating to each of the key topics described above

The non-systematic literature searches focused on high quality systematic reviews (of any level), recent clinical practice guidelines, and primary sources of evidence, where appropriate. The searches were conducted using EMBASE and Medline (using the EMBASE.com interface), the Cochrane Library, secondary HTA databases (e.g. NICE in the United Kingdom, CADTH in Canada), and guideline websites/databases (e.g. Guidelines International Network, National Guidelines Clearing House). Each search was restricted to studies that were published between 2002 and the literature search

date in September 2012. Cascade searching was undertaken to retrieve additional articles cited in key included papers and guidelines.

Similar to the systematic review, studies were only considered for inclusion in the non-systematic review of potential harms and other benefits of PSA testing if they met the pre-specified population, intervention and comparator (PIC) criteria outlined for the systematic review. Editorials, comments, book chapters, animal studies, correspondence, and news items were specifically excluded. Studies were also excluded if they were not reported in full (e.g. research or systematic review protocols, conference proceedings, articles published in abstract form) or were not published in the English language.

Key sources of evidence for this component of the project were critically appraised but formal quality assessment and data extraction forms were not completed for studies identified through non-systematic literature searches.

Papers discussing the adverse effects of investigations and treatments for prostate cancer were only considered relevant to the non-systematic review if it could be determined that the study population were screen-detected/asymptomatic men (rather than clinically diagnosed/symptomatic men).

References

NHMRC (2000). How to use the evidence: assessment and application of scientific evidence. National Health and Medical Research Council, Canberra ACT. Available at: http://www.nhmrc.gov.au/_files_nhmrc/file/publications/synopses/cp69.pdf

NHMRC (2007). Standards and procedures for externally developed guidelines. National Health and Medical Research Council, Canberra ACT. Available at: http://www.nhmrc.gov.au/_files_nhmrc/file/publications/synopses/nh56.pdf

NHMRC (2009). NHMRC additional levels of evidence and grades for recommendations for developers of guidelines. National Health and Medical Research Council, Canberra ACT. Available at: https://www.nhmrc.gov.au/_files_nhmrc/file/guidelines/developers/nhmrc_levels_grades_evidence_120423.pdf

Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002, 21(11):1539–58.

Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org

Appendix A Excluded studies

This appendix documents studies that met initial inclusion criteria for the systematic review but were excluded after full text review.

Studies excluded from the systematic review of Level I evidence		
Author (year)	**Description**	**Reason for exclusion**
Heidenreich et al (2012)	European Association of Urology evidence-based clinical practice guidelines on prostate cancer	Details of literature searches, quality assessments and critical appraisals not available
Loeb et al (2012)	Systematic review of the literature on baseline PSA testing for the prediction of prostate cancer risk and prognosis	The focus of the study was on evaluating the usefulness of obtaining baseline PSA testing from asymptomatic men at a young age (≤ 60 years); did not include relevant mortality/morbidity outcomes
Zhu et al (2012)	Systematic review of risk-based prostate cancer screening	The focus of the study was on identification of risk factors for prostate cancer; did not include relevant mortality/morbidity outcomes
Chou et al (2011a)	Summary of the 2011 AHRQ reviews of PSA testing and treatments for localised prostate cancer (Lin et al 2011; Chou et al 2011b): • Includes supplementary evidence on the benefits and harms of treatment of early-stage or screening detected prostate cancer	Mortality and morbidity results relevant to the systematic review are duplicated from 2011 AHRQ review of PSA testing (Lin et al 2011)
O'Rourke (2011)	Review of guidelines, recommendations and evidence regarding PSA screening	Not a systematic review
Brooks et al (2010)	American Cancer Society recommendations on screening for prostate cancer	Not a systematic review
Bryant and Hamdie (2008)	Systematic review of prostate cancer screening	Details of literature searches, quality assessments and critical appraisals not available
CADTH (2007)	Review of the effectiveness of PSA screening	Not a systematic review
Mambourg et al (2006)	Health technology assessment of PSA testing in Belgium	Details of literature searches, quality assessments and critical appraisals not in English
Mistry et al (2003)	Systematic review and meta-analysis of the diagnostic characteristics of PSA screening	The focus of the study was on sensitivity, specificity and positive predictive value; did not include relevant mortality/morbidity outcomes
Slaughter et al (2002)	Systematic review of PSA testing in asymptomatic men	The literature search strategies were limited to studies published between 1995-2001
Studies excluded from the Level II evidence update		
Author (year)	**Reason for exclusion**	
Boevee et al (2010)	Wrong outcomes	
Bul et al (2012)	Wrong outcomes	
Bul et al (2011)	Wrong population	
Carlsson et al (2011)	Wrong outcomes	
Author (year)	**Reason for exclusion**	
Djavan (2011)	Not an RCT	

Finne et al (2010)	Wrong outcomes
Hanley et al (2010)	Wrong outcomes
Lane et al (2010)	Ongoing RCT
Lujan et al (2012)	Sub-analysis of the Spanish arm of the ERSPC trial. Results are included in the overall ERSPC trial report.
Pinsky et al (2010)	Not an RCT
Sugimoto et al (2012)	Not an RCT
Van Leeuwen et al (2010)	Wrong outcomes
Van Leeuwen et al (2010)	Sub-analysis of the Rotterdam arm of the ERSPC trial. Results are included in the overall ERSPC trial report
Van Leeuwen et al (2012)	Wrong outcomes
Zhu et al (2011)	Sub-analysis of the Rotterdam arm of the ERSPC trial. Results are included in the overall ERSPC trial report
Zhu et al (2011)	Sub-analysis of the Rotterdam arm of the ERSPC trial. Results are included in the overall ERSPC trial report

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; CADTH, Canadian Agency for Drugs and Technologies in Health; PSA, prostate-specific antigen; RCT, randomised controlled trial.

Appendix B Included studies

Level I evidence

Study ID	Citation
Basch (2012)	Basch E, Oliver TK, Vickers A, Thompson I, Kantoff P, Parnes H, Loblaw DA, Roth B, Williams J, Nam RK. (2012) Screening for Prostate Cancer with Prostate-Specific Antigen Testing: American Society of Clinical Oncology Provisional Clinical Opinion. *J Clin Oncol* 30:3020-5.
Djulbegovic (2010)	Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, and Dahm P. (2010) Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. *BMJ* 341:c4543.
Hamashima (2009)	Hamashima C, Nakayama T, Sagawa M, Saito H, Sobue T. (2009) The Japanese guideline for prostate cancer screening. *Jpn J Clin Oncol* 39:339-51.
Ilic (2013) [Cochrane review]	Ilic D, Neuberger MM, Djulbegovic M, and Dahm P. (2013) Screening for prostate cancer. Cochrane Database of Systematic Reviews 2013 Issue 1. Art. No.: CD004720. DOI: 10.1002/14651858.CD004720.pub3.
Lin (2011) [AHRQ]	Lin K, Croswell JM, Koenig H, Lam C, and Maltz A. (2011) Prostate-specific antigen-based screening for prostate cancer: an evidence update for the U.S. preventive services task force. Evidence synthesis no. 90. AHRQ publication no. 12-05160-EF-1. Rockville, MD: Agency for Healthcare Research and Quality.
Lumen (2012)	Lumen N, Fonteyne V, de Meerleert G, Ost P, Villeirs G, Mottrie A, de Visschere P, de Troyer B, and Oosterlinck, W. (2012) Population screening for prostate cancer: an overview of available studies and meta-analysis. *Int J Urol* 19:100-108.
NZGG (2009)	New Zealand Guidelines Group. (2009) Cancer control strategy guidance completion: update of evidence for prostate-specific antigen (PSA) testing in asymptomatic men. Wellington: Ministry of Health.

Level II evidence

Study ID	Included citation(s), with primary publication marked in bold
PLCO	**Andriole GL, Crawford DE, Grubb III RL, Buys SS, Chia D, Church TR, Fouad MN, Isaacs C et al. (2012) Prostate cancer screening in the randomised prostate, lung, colorectal, and ovarian cancer screening trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst* 104:125-132.**
	Andriole GL, Crawford DE, Grub III RL, Buys, SS, Chia D, Church TR, Fouad MN, Gelmann EP et al. (2009) Mortality results from a randomised prostate cancer screening trial. *N Eng J Med* 360:1310-1319.
	Crawford ED, Grubb R, Black A, Andriole GL, Chen MH, Izmirlian G, Berg CD, and D'Amico AV. (2011) Comorbidity and mortality results from a randomised prostate cancer screening trial. *J Clin Oncol* 29:355-361.
	Gulati R, Tsodikov A, Wever EM, Mariotto AB, Heijnsdijk AM, Katcher J, de Koning HJ, and Etzioni R. (2012) The impact of PLCO control arm contamination on perceived PSA screening efficacy. *Cancer Cause Control* 23:827-835.
	Johnson DB. (2006) The effects of an abnormal cancer screening test on health related quality of life. *Intl J Cancer Res* 2(3): 277-289.
	Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, Fogel R, Gelmann EP et al. (2000) Design of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials* 21(6 Suppl): 273S-309S.
	Taylor KL, Luta G, Miller AB, Church TR, Kelly SP, Muenz LR, Davis KM, Dawson DL et al. (2012) Long-term disease-specific functioning among prostate cancer survivors and noncancer controls in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *J Clin Oncol* 30(22): 2768-2775.
	Taylor KL, Shelby R, Gelmann E, and McGuire C. (2004) Quality of life and trial adherence among participants in the Prostate, lung, Colorectal and Ovarian Cancer Screening Trial. *J Natl Cancer Inst* 96(14): 1083-1094.

ERSPC	Carlsson SV, Holmberg E, Moss SM, Roobol MJ, Schroder FH, Tammela TLJ, Aus G, Auvinen AP, and Hugosson J. (2011) No excess mortality after prostate biopsy: results from the European Randomised Study of Screening for Prostate Cancer. *BJU Int* 107:1912-1917.
	Heijnsdijk EAM, Wever EM, Auvinen A, Hugosson J, Ciatto S, Nelen V, Kwiatkowski M, Villers A et al. (2012) Quality of life effects of prostate specific antigen screening. *N Eng J Med* 367(7): 595-605.
	Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schroder FJ, and Steyerberg EW. (2010) Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomised screening trial (ERSPC section Rotterdam). *Int J Cancer* 127:2639-2644.
	Kilpelainen TP, Tammela TLJ, Roobol M, Hugosson J, Ciatto S, Nelen V, Moss S, Maattaenen L, and Auvinen A. (2011) False-positive screening results in the European randomised study of screening for prostate cancer. *Eur J Cancer* 47:2698-2705.
	Kilpelainen TP, Auvinen A, Maattanen L, Kujala P, Ruutu M, Stenman UH, and Tammela TLJ. (2010) Results of the three rounds of the Finnish prostate cancer screening trial – the incidence of advanced cancer is decreased by screening. *Int J Cancer* 127(7):1699-1705.
	Roobol M, Kerkhoff M, Schroder F, Cuzik J, Sasieni P, Hakama M, et al. (2009) Prostate cancer mortality reduction by prostate-specific antigen based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 56:585–91.
	Schroder FH, Hugosson J, Carlsson S, Tammela T, Maattanen L, Auvinen A, Kwiatkowski M, Recker F et al. (2012) Screening for prostate cancer decreases the risk of developing metatastic disease: findings from the European Rnadomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 62(5):745-752.
	Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M et al. (2012) Prostate cancer mortality at 11 years of follow-up. *N Eng J Med* 366:981-990.
	Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H et al. (2009) Screening and prostate cancer mortality in a randomized European study. *N Eng J Med* 360:1320-1328.
Goteborg	Bergdahl A, Aus G, Lilja H, and Hugosson J. (2009). Risk of dying from prostate cancer in men randomised to screening. *Cancer* 115(24): 5672-5679.
	Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, Pihl CG, Stranne J et al. (2010) Mortality results from the Goteborg randomised population-based prostate cancer screening trial. *Lancet Oncol* 11:725-32.
Norrkoping	**Sandblom G, Varenhorst E, Rosell J, Lofman O and Carlsson P. (2011) Randomised prostate cancer screening trial: 20 year follow-up. *BMJ* 342: d1539.**
	Sandblom G, Varenhorst E, Lofman O, Rosell J, Carlsson P. (2004) Clinical consequences of screening for prostate cancer: 15 years follow-up of a randomised controlled trial in Sweden. *Eur Urol* 46:717-724.
Stockholm	**Kjellman A, Akre O, Norming U, Tornblom M, and Gustafsson O. (2009) 15-year followup of a population based prostate cancer screening study. *J Urol* 181:1615-1621.**
Quebec	**Labrie F, Candas B, Cusan L, Gomez JL, Belanger A, Brousseau g, Chebrette E and Levesque J. (2004) Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. *Prostate* 59(3): 311-318.**

Appendix C Sources of funding and declared interests of the authors in each included Level I study

Basch et al (2012)	
Sources of funding	None reported
Declared interests of the authors	All authors completed the disclosure declaration. The following author(s) and/or an author's immediate family member indicated a financial or other interest that is relevant to the subject matter under consideration in this article: Employment or leadership position: none Consultant or advisory role: Andrew Vickers, Genomic Health (compensated), GlaxoSmithKline (compensated) Stock ownership: none Honoraria: Andrew Vickers, Genomic Health, GlaxoSmithKline Research funding: none Expert testimony: none Other remuneration: none
Djulbegovic et al (2010)	
Sources of funding	Department of Urology, University of Florida. Dennis W Jahnigen Career Development Scholars Award through the American Geriatrics Society *"The funders had no role in study design, the collection, analysis and interpretation of data, the writing of the report or the decision to submit the article for publication."*
Declared interests of the authors	All authors completed the Unified Competing Interest form and declared no support from any institution for the submitted work; no financial relationships with any institutions that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.
Hamashima et al (2009)	
Sources of funding	Grant-in-Aid for Cancer Control from Ministry of Health, Labour and Welfare of Japan (Grant number 15-3)
Declared interests of the authors	None declared
Ilic et al (2013) [Cochrane review]	
Sources of funding	Department of Urology, College of Medicine, University of Florida, USA Malcolm Randall Veterans Affairs Medical Centre, Gainesville, Florida, USA Dennis W. Jahnigen Career Development Scholars Award by the American Geriatrics Society, USA
Declared interests of the authors	None known
Lin et al (2012) [AHRQ]	
Sources of funding	Agency for Healthcare Research and Quality. This review did not receive specific funding.
Declared interests of the authors	Not reported
Lumen et al (2012)	
Sources of funding	Ghent University Hospital, Belgium
Declared interests of the authors	None declared
NZGG (2009)	
Sources of funding	Ministry of Health, New Zealand
Declared interests of the authors	Not reported

www.nhmrc.gov.au
GPO Box 1421, Canberra ACT 2601
16 Marcus Clarke Street, Canberra City ACT
T. 13 000 NHMRC (13 000 64672) or +61 2 6217 9000 F. 61 2 6217 9100 E. nhmrc@nhmrc.gov.au