Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials

Mia Djulbegovic, student,[1] Rebecca J Beyth, associate professor,[2] Molly M Neuberger, research assistant,[1] Taryn L Stoffs, research assistant,[1] Johannes Vieweg, professor and chairman,[1] Benjamin Djulbegovic, professor,[3] Philipp Dahm, associate professor[1]

[1]Department of Urology and Prostate Disease Center, University of Florida, College of Medicine, PO Box 100247, Gainesville, Florida 32610-0247, USA

[2]Department of Medicine, University of Florida, College of Medicine, PO Box 100277, Gainesville, Florida

[3]Center and Division for Evidence Based Medicine and Outcomes Research, University of South Florida, MDC 27, Tampa, Florida 33612

Correspondence to: P Dahm
p.dahm@urology.ufl.edu

ABSTRACT

Objective To examine the evidence on the benefits and harms of screening for prostate cancer.

Design Systematic review and meta-analysis of randomised controlled trials.

Data sources Electronic databases including Medline, Embase, CENTRAL, abstract proceedings, and reference lists up to July 2010.

Review methods Included studies were randomised controlled trials comparing screening by prostate specific antigen with or without digital rectal examination versus no screening. Data abstraction and assessment of methodological quality with the GRADE approach was assessed by two independent reviewers and verified by the primary investigator. Mantel-Haenszel and inverse variance estimates were calculated and pooled under a random effects model expressing data as relative risks and 95% confidence intervals.

Results Six randomised controlled trials with a total of 387 286 participants that met inclusion criteria were analysed. Screening was associated with an increased probability of receiving a diagnosis of prostate cancer (relative risk 1.46, 95% confidence interval 1.21 to 1.77; P<0.001) and stage I prostate cancer (1.95, 1.22 to 3.13; P=0.005). There was no significant effect of screening on death from prostate cancer (0.88, 0.71 to 1.09; P=0.25) or overall mortality (0.99, 0.97 to 1.01; P=0.44). All trials had one or more substantial methodological limitations. None provided data on the effects of screening on participants' quality of life. Little information was provided about potential harms associated with screening.

Conclusions The existing evidence from randomised controlled trials does not support the routine use of screening for prostate cancer with prostate specific antigen with or without digital rectal examination.

INTRODUCTION

Prostate cancer is the most common non-skin cancer among men worldwide[1] and, after lung cancer, is the second leading cause of deaths from cancer in men in the United States.[2] Screening has been advocated as a means of detecting prostate cancer in the early stages, which are amenable to local interventions with curative intent, to decrease overall and disease specific mortality.[3] The benefits and harms of prostate cancer screening, however, have become the topic of controversy, as reflected by numerous recent editorials,[4-7] position statements, and guidance documents.[8-10]

Population based recommendations for cancer screening should ideally be based on high quality evidence derived from systematic reviews of randomised controlled trials that document a positive impact of screening on outcomes that are the most important to patients.[11] In 2006, a systematic review published in the Cochrane Library concluded that there was insufficient evidence to either support or refute the routine use of mass, selective, or opportunistic screening compared with no screening.[12] This Cochrane systematic review was based on two randomised controlled trials that enrolled 55 512 participants overall but was limited by substantial methodological weaknesses in the design, conduct, and analysis of the included studies. The evidence drawn from this systematic review did not show that screening improved outcomes. By 2010, four additional trials[4 13-15] enrolling 351 531 participants had been published, thereby providing strong impetus for an updated synthesis of research evidence.

We performed a systematic review and meta-analysis on the role of screening for prostate cancer to guide decision making in health policy. Specifically, we assessed the question of whether in men without a previous history of prostate cancer, screening by testing for prostate specific antigen with or without digital rectal examination when compared with no screening affects the two most important outcomes to patients: overall and disease specific mortality.

METHODS

Data sources and searches

We conducted a systematic search of electronic databases, abstract proceedings of major scientific meetings, and bibliographies of all eligible studies from 1 January 2005 to the present (the last systematic search was dated 13 July 2010) to identify all relevant studies since the comprehensive search conducted for the systematic review published in the Cochrane Library in 2006.[12] Electronic databases searched included

Medline (PubMed), Embase, and the Cochrane Registry of Controlled Trials (CENTRAL). The search strategy involved combining a methodological filter to identify randomised controlled trials[16] with subject specific terms related to screening for prostate cancer (("Mass Screening"[Mesh] OR "Early Detection of Cancer"[Mesh]) AND "Prostatic Neoplasms"[Mesh]).

The manual search included abstracts presented at the American Urological Association (AUA), European Association of Urology (EAU), and American Society of Clinical Oncology (ASCO) meetings from 2005 to 2010. We also searched for additional systematic reviews and narrative reviews on the topic to identify eligible trials. Studies were considered irrespective of language or publication status. Two independent reviewers (MD and MMN) performed all aspects of the search strategy, examined the abstracts of all citations for relevance to our predefined inclusion criteria, and reviewed the full text articles in detail as indicated. PD reviewed and arbitrated any disagreements.

Study selection

Randomised controlled trials comparing screening of asymptomatic men for prostate cancer versus no screening were eligible for inclusion. The screening intervention was defined as testing for prostate specific antigen with or without digital rectal examination. We did not include trials with participants with previously diagnosed prostate cancer.

Data extraction and quality assessment

A standardised form was created, piloted, and then used to abstract the available data for the predefined outcomes of interest. These were: all cause mortality and death from prostate cancer, diagnosis of prostate cancer, effect of screening on stage at diagnosis, false positive and false negative results, harms of screening, quality of life, and cost effectiveness. We used the 2010 American Joint Committee on Cancer system for prostate cancer staging.[17] Two authors (MD and MMN) independently extracted data. Disagreements were resolved by discussion, consensus, and arbitration by a third author (PD). Data were extracted on the methodological domains relevant to minimising bias and random error in the analysis of trials by using the Cochrane methods for assessing risk of bias[18] and GRADE methods.[19] Specifically, we assessed study limitations by evaluating the method of randomisation, allocation concealment, blinding, analysis by intention to screen, contamination of the control arm, and completeness of follow-up. As per GRADE,[19] we further assessed the quality of evidence with regard to inconsistency (heterogeneity), indirectness, imprecision, and other potential sources of bias, such as publication and reporting bias (see below). GRADE criteria were then applied to downgrade the quality of evidence when indicated on an outcome specific basis. The quality of evidence for an individual outcome was ultimately rated as high, moderate, low, or very low. The review protocol is available from the authors on request.

Data synthesis and analysis

Relative risks were used to summarise the effect of screening intervention for all outcomes. Mantel-Haenszel estimates were calculated based on the number of events per 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, we used the inverse variance method. Heterogeneity was assessed by examining clinical characteristics of included studies as well as by formal statistical testing with χ^2 and I^2.[18] The possibility of publication bias was assessed with Begg and Egger funnel plots.[20] The results of these tests are not separately reported, however, because this method is known to be unreliable when there are fewer than 10 studies in the meta-analysis[18] and because our qualitative analysis indicated a high likelihood of reporting bias. Meta-analysis was performed with RevMan 5[21] according to the PRISMA guidelines.[22] We used the GRADE method to summarise findings.[19] We also carried out predefined subgroup analyses for participants' age and stage at diagnosis and sensitivity analyses based on methodological quality parameters.

RESULTS

Systematic review

The systematic literature search identified 493 relevant references (fig 1). After screening titles and abstracts, we excluded 463 non-relevant articles. The 30 remaining articles were retrieved in full text for formal review. After independent review, 24 studies were excluded because they were subset or secondary analyses of the same trials. We also retrieved recent review articles,[23] position papers and editorials,[5 24] and a previously published systematic review,[12] but these did not yield any additional relevant trials.

Table 1 gives details of the six identified studies that met predefined inclusion and exclusion criteria. These
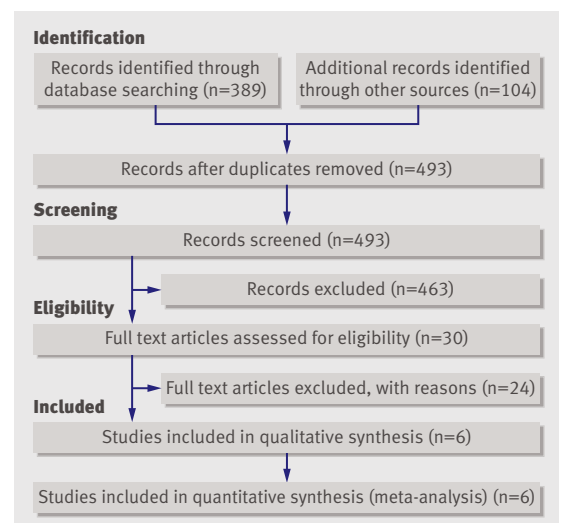

Fig 1 | Search strategy to identify trials of screening for prostate cancer

studies, four of which were published since 2009,[4 13-15] enrolled a total of 387 286 participants randomised to either prostate cancer screening or no screening. These trials included the European randomised study of screening for prostate cancer (ERSPC),[4] the prostate, lung, colorectal, and ovarian (PLCO) screening trial,[14 25] the French ERSPC,[13] which was part of the original ERSPC[4] but reported separately, and the Gothenburg trial,[15] which included participants previously reported in the ERSPC trial. The Gothenburg trial[15] enrolled 19 904 participants in three birth cohorts (1930-4, 1935-9, and 1940-4), of which two (1930-4, 1935-9) were included in the ERSPC publication[4] as the Swedish cohort (n=11 852). Only the participants of the 1940-4 birth cohort (n=8057) of the Gothenburg trial[15] represented study participants not previously reported on. For participants of the 1930-4 (n=5563) and 1935-9 (n=6284) birth cohorts, the Gothenburg trial[15] provided longer follow-up than previously reported in the ERSPC.[4] We included these data in the analyses by excluding the corresponding participants from the ERSPC.[4]

All but one study included measurement of prostate specific antigen as a screening test in all participants; the Norrkoping study[26 27] initially used only digital rectal examination but then used a combination of prostate specific antigen and digital rectal examination. Three of the six studies did not consistently use digital rectal examination in all participants; in the ERSPC the screening method differed by participating country and was mostly based on prostate specific antigen.[4] In the French ERSPC, only prostate specific antigen testing, not digital rectal examination, was used.[13] Finally, in the Gothenburg study screening was based on

prostate specific antigen testing alone, and participants received a digital rectal examination only if the test result was abnormal.[15]

Four studies provided information on all cause mortality,[4 14 15 26 27] five studies on deaths from prostate cancer,[4 14 15 26-30] and five studies on diagnosis of prostate cancer.[4 13-15 26 27] Length of follow-up ranged from about four to 15 years.[13 26] All but the Quebec study[28-30] and the Gothenburg study[15] provided usable information on cancer stage at diagnosis. The ERSPC study[4] and, to a limited extent, the Gothenburg study[15] allowed subgroup analyses for death from prostate cancer based on age groups, but only the Gothenburg study provided age specific information for all cause mortality.[15] There were no major discrepancies between reviewers with regard to trial inclusion or data extraction. Minor discrepancies in extraction of clinical characteristics were resolved by consensus and arbitration by a third independent reviewer (PD).

Assessment of quality of evidence

Common study limitations included inadequate randomisation and allocation concealment, non-reporting of study withdrawals and participants lost to follow-up, lack of blinding of outcome assessor, contamination, and failure to perform an intention to treat analysis (table 1). These prompted downgrading of the overall quality of evidence for all individual outcomes (table 2). Failure of studies that reported on death from prostate cancer to provide information on all cause mortality raised further concerns about reporting bias. The evidence on diagnosis of prostate cancer overall and diagnosis of stages I and II prostate cancer was downgraded for inconsistency. For the diagnosis of stage II prostate

Table 1 | Characteristics and methodological quality of randomised controlled trials of screening for prostate cancer

Study	No of participants randomised		Age range (years)	Screening test	Random-isation	Allocation concealment	Description of loss to follow-up	Blinding of outcome assessors	Contamination (of control group)	Intention to screen analysis	Approximate median follow-up time (years)
	Screened group	Control group									
Quebec[28-30]	31 133	15 353	45-80	PSA + DRE	Adequate	Not adequate*†	No	Not adequate‡	Not provided§	Not analysed but data provided	11
Norrkoping[26 27]	1494	7532	50-69	DRE initially; PSA + DRE after 1993	Quasi-random-isation**	Not adequate*	Yes	Not adequate‡	Not provided††	Yes	13 (diagnosis), 15 (death)
ERSPC[4]	72 890‡‡	89 353‡‡	55-69	PSA + DRE	Adequate	Not adequate*†	Unclear§§	Yes	Yes (20%)	Yes	9
French ERSPC[13]	42 590	42 191	55-69	PSA	Unclear***	Unclear*	Unclear§§	Unclear***	Unclear***	Unclear***	4
PLCO[14]	38 343	38 350	55-74	PSA+DRE	Adequate	Adequate	Yes	Yes	Yes (40-52%)	Yes	11.5
Gothenburg[15]	9952	9952	50-64	PSA	Adequate	Not Adequate†	Yes	Yes	"Low," details not provided	Yes	14

PSA=prostate specific antigen.
DRE=digital rectal examination.
*Not reported/could not be assessed (adequacy of allocation concealment could not be deduced from available reports).
†Participants randomised to screening versus control before formal study enrolment, thus raising possibility of selection bias.
‡Cause of death determined based on registry data.
§Study reported that contamination rate could not be assessed.
**Every sixth man from list of dates of birth assigned to screening.
††Contamination rate with respect to control group not provided.
‡‡Number of participants includes 1930-4 and 1935-9 birth cohorts (age 60-64 and 55-59 at randomisation, respectively) of Gothenburg study.
§§No data provided on completeness of follow-up or could not be assessed.
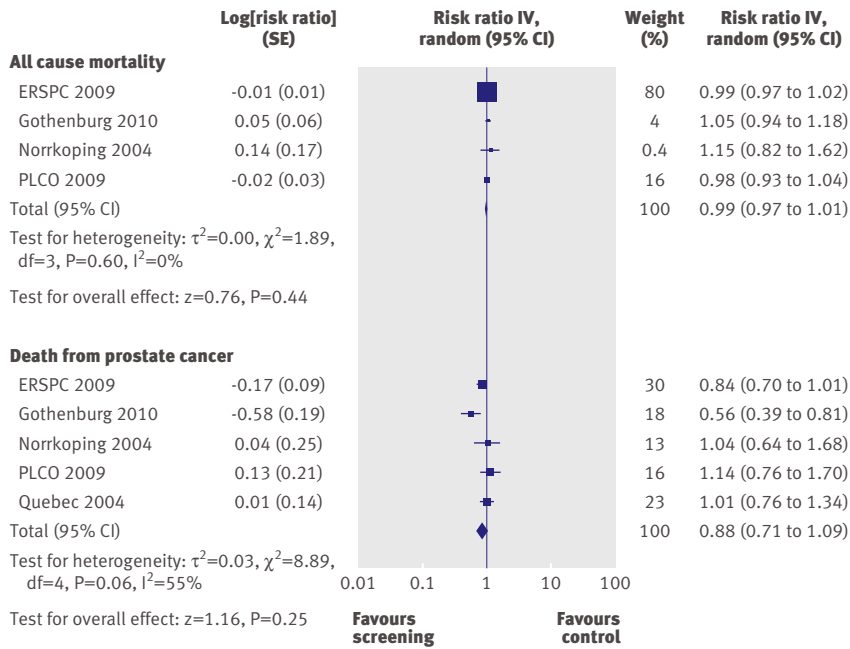***Could not be assessed or not reported.

	Log[risk ratio] (SE)	Risk ratio IV, random (95% CI)	Weight (%)	Risk ratio IV, random (95% CI)
All cause mortality				
ERSPC 2009	-0.01 (0.01)		80	0.99 (0.97 to 1.02)
Gothenburg 2010	0.05 (0.06)		4	1.05 (0.94 to 1.18)
Norrkoping 2004	0.14 (0.17)		0.4	1.15 (0.82 to 1.62)
PLCO 2009	-0.02 (0.03)		16	0.98 (0.93 to 1.04)
Total (95% CI)			100	0.99 (0.97 to 1.01)

Test for heterogeneity: $\tau^2=0.00$, $\chi^2=1.89$, df=3, P=0.60, $I^2=0\%$

Test for overall effect: z=0.76, P=0.44

Death from prostate cancer				
ERSPC 2009	-0.17 (0.09)		30	0.84 (0.70 to 1.01)
Gothenburg 2010	-0.58 (0.19)		18	0.56 (0.39 to 0.81)
Norrkoping 2004	0.04 (0.25)		13	1.04 (0.64 to 1.68)
PLCO 2009	0.13 (0.21)		16	1.14 (0.76 to 1.70)
Quebec 2004	0.01 (0.14)		23	1.01 (0.76 to 1.34)
Total (95% CI)			100	0.88 (0.71 to 1.09)

Test for heterogeneity: $\tau^2=0.03$, $\chi^2=8.89$, df=4, P=0.06, $I^2=55\%$

Test for overall effect: z=1.16, P=0.25

0.01 0.1 1 10 100

Favours screening — Favours control

Fig 2 | Effects of screening on all cause mortality and death from prostate cancer

cancer, there was a wide confidence interval for the effect size, which included both appreciable benefit and harm and led to further downgrading for imprecision. Overall, the quality of evidence was rated as moderate for both all cause mortality and death from prostate cancer (table 2).[19]

Quantitative assessment

All cause mortality

Four trials that included 256 019 randomised participants contributed information on all cause mortality.[4 14 15 26 27] As event rates were not available in all studies, we used the inverse variance method to pool data from individual trials, resulting in a relative risk of 0.99 (95% confidence interval 0.97 to 1.01; P=0.44; fig 2). There was no significant heterogeneity among these trials ($I^2=0\%$, $\chi^2=1.89$; P=0.60). As it is not

plausible that authors would report data on cause specific mortality without having collected data on overall mortality we suspect outcome reporting bias[19] for the Quebec trial.[28-30]

Death from prostate cancer

Data on deaths from prostate cancer were available from five randomised controlled trials.[4 14 15 26 28] The analysis included 302 500 randomised participants. With the inverse variance method, the calculated relative risk was 0.88 (0.71 to 1.09; P=0.25) when analysed in an intention to screen analysis (fig 2). There was considerable heterogeneity among these trials ($I^2=55\%$, $\chi^2=8.89$; P=0.06).

Prostate cancer diagnosis

Five trials contributed information on diagnosis of prostate cancer in 340 800 randomised participants.[4 13-15 26 27] The Quebec study did not report on disease stages in the control arm. There were 10 328 men with a diagnosis of prostate cancer among the 159 372 men enrolled in the screening group compared with 7968 in the 181 428 controls, resulting in a relative risk of 1.46 (1.21 to 1.77; P<0.001; fig 3) in favour of screening. There was a high degree of heterogeneity in these trials ($I^2=97\%$, $\chi^2=126.69$; P<0.001).

The subgroup analysis for stage I prostate cancer was based on 3789 men with a diagnosis of stage I prostate cancer among the 155 317 men in the screening group compared with 1971 in the 177 426 control group, resulting in a relative risk of 1.95 (1.22 to 3.13; P=0.005) in favour of screening (fig 4). There was a high degree of heterogeneity ($I^2=96\%$, $\chi^2=79.32$; P<0.001).

Stage II prostate cancer was diagnosed in 5114 of the 155 317 men in the screening group and 4035 of the 177 426 controls, resulting in a relative risk of 1.39 (0.99 to 1.95; P=0.05; fig 4). In this analysis, there was significant heterogeneity ($I^2=97\%$, $\chi^2=114.38$; P<0.001).

Table 2 | Summary of findings in trials on screening for prostate cancer

	Illustrative comparative risks* (95% CI)				
Outcomes	Event rate (per 1000) without screening	Event rate (per 1000) with screening	Relative risk (95% CI)	No of participants	Quality of evidence (GRADE)†
All cause mortality (inverse variance)	200‡	198 (194 to 202)	0.99 (0.97 to 1.01)	256 019[4 14 15 26]	Moderate
Death from prostate cancer (inverse variance)	8‡	7 (6 to 9)	0.88 (0.71 to 1.09)	302 500[4 14 15 26 28]	Moderate
Prostate cancer diagnosis	44	64 (53 to 78)	1.46 (1.21 to 1.77)	340 800[4 13-15 26]	Low
Effects of screening on stage:					
Stage I	11	21 (13 to 34)	1.95 (1.22 to 3.13)	332 743[4 13 14 26]	Low
Stage II	23	32 (23 to 45)	1.39 (0.99 to 1.95)	332 743[4 13 14 26]	Very low
Stages III-IV	5	5 (4 to 5)	0.94 (0.85 to 1.04)	332 743[4 13 14 26]	Moderate

*Event rate in no screening arm is mean risk in control groups across studies. Event rate in screening arm (and its 95% confidence interval) is based on assumed risk in comparison group and relative effect (relative risk) of intervention (and its 95% CI).

†High quality=further research is very unlikely to change our confidence in the estimate of effect; moderate quality=further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; low quality=further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; very low quality=we are very uncertain about the estimate.

‡Event rates not available in all trials; representative rates in control arms are based on Gothenburg trial; inverse variance method used to pool data from individual trials.
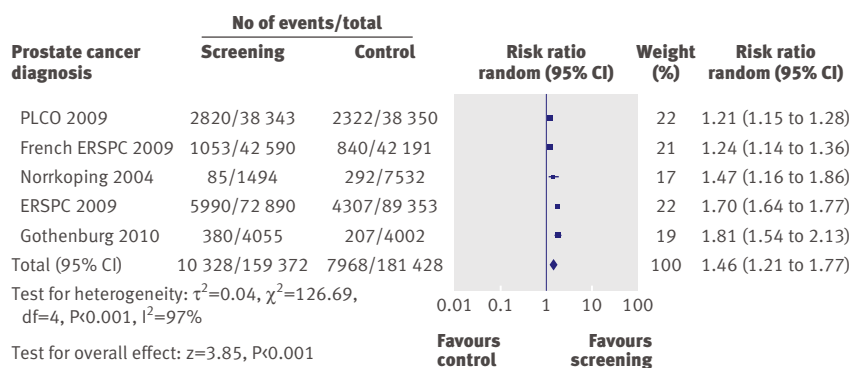
Prostate cancer diagnosis	No of events/total		Risk ratio random (95% CI)	Weight (%)	Risk ratio random (95% CI)
	Screening	Control			
PLCO 2009	2820/38 343	2322/38 350		22	1.21 (1.15 to 1.28)
French ERSPC 2009	1053/42 590	840/42 191		21	1.24 (1.14 to 1.36)
Norrkoping 2004	85/1494	292/7532		17	1.47 (1.16 to 1.86)
ERSPC 2009	5990/72 890	4307/89 353		22	1.70 (1.64 to 1.77)
Gothenburg 2010	380/4055	207/4002		19	1.81 (1.54 to 2.13)
Total (95% CI)	10 328/159 372	7968/181 428		100	1.46 (1.21 to 1.77)

Test for heterogeneity: $\tau^2=0.04$, $\chi^2=126.69$, df=4, P<0.001, $I^2=97\%$

Test for overall effect: z=3.85, P<0.001

0.01 0.1 1 10 100

Favours control   Favours screening

Fig 3 | Effects of screening on diagnosis of prostate cancer

Data on the detection of stages III and IV prostate cancer were based on 332 743 randomised participants. Stages III and IV cancer were diagnosed in 701 of 155 317 men enrolled in the screening group and 975 of 177 426 controls, resulting in a relative risk of 0.94 (0.85 to 1.04; P=0.22; fig 4). There was no significant heterogeneity in this analysis ($I^2=0\%$, $\chi^2=1.22$; P=0.75).

Subgroup analysis based on age
Age specific information for all cause mortality was limited to the Gothenburg study,[15] which reported

	No of events/total		Risk ratio random (95% CI)	Weight (%)	Risk ratio random (95% CI)
	Screening	Control			
Stage I					
ERSPC 2009	3276/72 890	1553/89 353		29	2.59 (2.44 to 2.74)
French ERSPC 2009	455/42 590	339/42 191		29	1.33 (1.16 to 1.53)
Norrkoping 2004	40/1494	64/7532		24	3.15 (2.13 to 4.66)
PLCO 2009	18/38 343	15/38 350		18	1.20 (0.60 to 2.38)
Total (95% CI)	3789/155 317	1971/177 426		100	1.95 (1.22 to 3.13)

Test for heterogeneity: $\tau^2=0.20$, $\chi^2=79.32$, df=3, P<0.001, $I^2=96\%$

Test for overall effect: z=2.80, P=0.005

Stage II					
ERSPC 2009	1571/72 890	984/89 353		28	1.96 (1.81 to 2.12)
French ERSPC 2009	225/42 590	169/42 191		26	1.32 (1.08 to 1.61)
Norrkoping 2004	21/1494	92/7532		18	1.15 (0.72 to 1.84)
PLCO 2009	3297/38 343	2790/38 350		28	1.18 (1.13 to 1.24)
Total (95% CI)	5114/155 317	4035/177 426		100	1.39 (0.99 to 1.95)

Test for heterogeneity: $\tau^2=0.10$, $\chi^2=114.38$, df=3, P<0.001, $I^2=97\%$

Test for overall effect: z=1.93, P=0.05

Stage III and IV					
ERSPC 2009	516/72 890	676/89 353		74	0.94 (0.83 to 1.05)
French ERSPC 2009	39/42 590	32/42 191		4	1.21 (0.76 to 1.93)
Norrkoping 2004	24/1494	132/7532		5	0.92 (0.60 to 1.41)
PLCO 2009	122/38 343	135/38 350		16	0.90 (0.71 to 1.15)
Total (95% CI)	701/155 317	975/177 426		100	0.94 (0.85 to 1.04)

Test for heterogeneity: $\tau^2=0.00$, $\chi^2=1.22$, df=3, P=0.75, $I^2=0\%$

0.01 0.1 1 10 100

Favours control   Favours screening
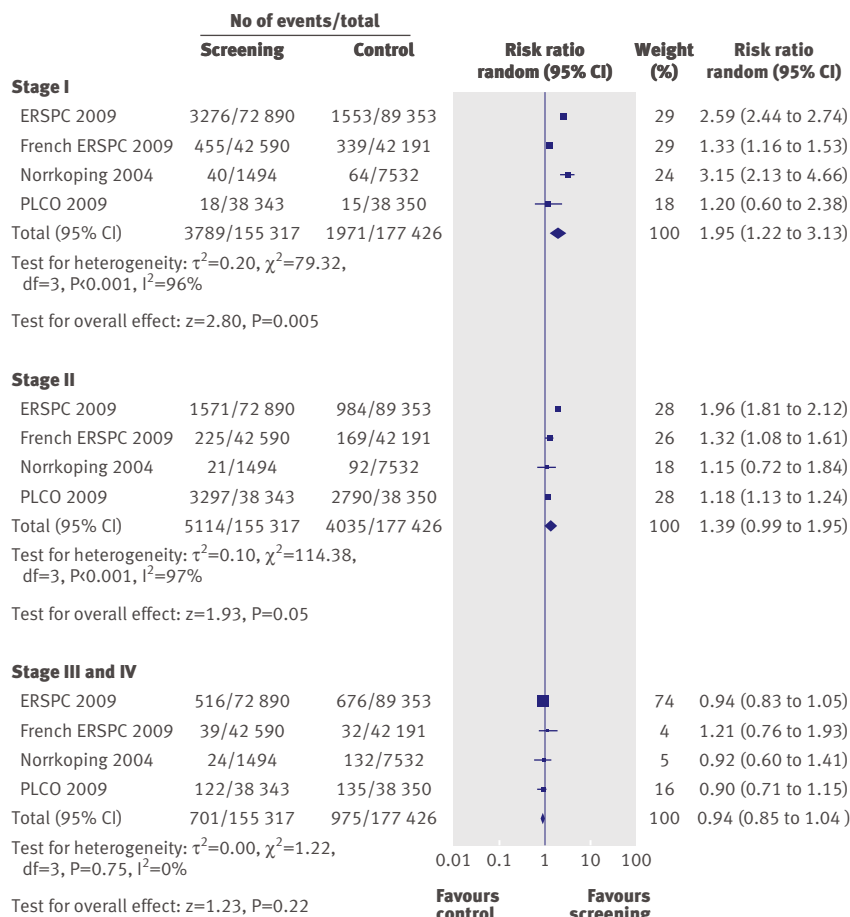
Test for overall effect: z=1.23, P=0.22

Fig 4 | Effects of screening on prostate cancer stage at time of diagnosis

relative risks of 1.05 (0.94 to 1.18), 0.99 (0.90 to 1.09), and 0.99 (0.91 to 1.07) for participants aged 50-54, 55-59, and 60-64, respectively. Data on the effect of screening on death from prostate cancer were largely limited to the ERSPC,[4] with additional information from the Gothenburg study[15] only for men aged 50-54. The relative risks of death from prostate cancer in the screening compared with the control arms for participants aged 50-54, 55-59, 60-64, 65-69, and 70-74 were 0.90 (0.39 to 2.10), 0.73 (0.53 to 1.00), 0.94 (0.69 to 1.27), 0.74 (0.56 to 0.99), and 1.26 (0.80 to 1.99), respectively.

Limited information on age specific diagnosis of prostate cancer based on digital rectal examination alone was available from the first two screening rounds of the Norrkoping study.[27] In addition, the Gothenburg study contributed data based on prostate specific antigen testing for men aged 50-54, 55-59, and 60-64.[15] The relative risks for a diagnosis of prostate cancer in the screening compared with the control arms for participants aged 50-54, 55-59, 60-64, and 65-69 were 1.81 (1.53 to 2.13), 1.62 (1.40 to 1.88), 1.38 (1.19 to 1.61) and 2.44 (1.41 to 4.25), respectively.

Other outcomes
The randomised controlled trials we included failed to report complications rates in the screening and control group so we could not quantitatively pool data. The ERSPC trial did not include updates on the adverse events it reported in 2002.[31 32] A recent abstract based on three ERSPC centres reported no excess mortality associated with prostate biopsies in the screening arm.[33] The PLCO trial reported that digital rectal examination led to bleeding or pain at a rate of 0.3 per 10 000 screenings, and the prostate specific antigen test included three episodes of fainting per 10 000 screenings.[14] Medical complications (such as infections, bleeding, clot formation, and urinary difficulties) occurred in 68 per 10 000 diagnostic evaluations. No other studies reported data in either a qualitative or quantitative format. The Norrkoping study[26] and the ERSPC[4] reported false positive rates with screening of 82.5% and 75.9%, respectively, as verified by subsequent prostate biopsy. 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.

The Norrkoping study reported that screening costs in the 1990s were £1640 per detected cancer and £2343 per detected and cured cancer.[27] In light of our findings indicating no effect of screening on survival—that is, improvement in cure rates—these cost effectiveness results do not seem plausible. No other study reported data on costs or cost effectiveness.

DISCUSSION
Main findings
In this systematic review and meta-analysis of prostate cancer screening we failed to find a significant impact of prostate cancer screening on overall mortality or death from prostate cancer, the most critical outcomes

for patients. Evidence for both all cause mortality and death from prostate cancer was of moderate quality according to the GRADE approach. In contrast, based on low quality evidence, screening was associated with a 46% relative increase in diagnoses of prostate cancer in the screening arm compared with no screening. A predefined subgroup analysis based on disease stage indicated that this relative increase was attributable mainly to an increase in the number of men diagnosed with stage I prostate cancer. There was no significant impact of screening on the diagnosis of stage II and stages III and IV prostate cancer. These findings suggest that screening leads to an increase in diagnosis of early stage prostate cancer that does not seem to translate into a benefit in overall survival and survival specific to prostate cancer. On average, 20 more men will be detected with prostate cancer (95% confidence interval 9 to 34) per 1000 patients screened. The finding that an increased rate of diagnosis fails to translate into improved overall and disease specific mortality rates is probably multifactorial and relates in part to the prolonged and relatively slow natural course in many patients, particularly those with low grade prostate cancer.[34 35] Our results confirm previously voiced concerns about overdiagnosis of prostate cancer[6 7 23 24 36]—that is, detection of cancer that will not negatively affect survival. These findings are supported by studies that have used statistical modelling of population based data from a tumour registry[37] and the ERSPC study[38 39] and found overdiagnosis rates of 29% and 56%, respectively.

Strengths and limitations

We included the new evidence provided by the publication of the PLCO,[14] ERSPC,[4] and French ERSPC[13] as well as the Gothenburg trial.[15] These recently published randomised controlled trials conducted in the US and Europe included over 350 000 participants, representing major international efforts to address the controversy surrounding prostate cancer screening. Our systematic review, which includes these studies, therefore reflects the current best evidence on which decision making on health policy should be based.

We rated the methodological quality and risk of bias in studies meeting predefined inclusion and exclusion criteria using the GRADE approach. Limitations of our study relate to the quality and quantity of the available evidence on this topic. Unearthing the limitations in the evidence base, however, is as important as the overall findings related to the effects of screening. Considering the tremendous impact of decisions for or against prostate cancer screening on health policy, it is surprising that few randomised controlled trials have examined this issue with sufficient rigor. Existing studies have considerable methodological shortcomings that resulted in downgrading of the quality of evidence from high to moderate for all cause mortality and disease specific mortality (critical outcomes) and from high to low quality evidence for diagnosis of prostate cancer (important outcome), as well as considerable inconsistency. Most of the potential biases identified in the individual trials (such as lack of allocation concealment or intention to screen analysis) would be expected to favour the screening arm. On the other hand, contamination of the non-screening arm, a possible issue in all studies and one that was explicitly reported as being a major issue in the PLCO study, potentially introduced a bias towards not finding a benefit of screening.[23] Another limitation of the available evidence relates to the short length of follow-up of reported studies. Assuming an estimated lead time bias of five and a half to seven years, follow-up might not have been long enough to detect differences in mortality given the low number of deaths from prostate cancer.[39] Lastly, there was 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.

One of our predefined objectives was the analysis of the effect of screening interventions based on participants' age. This analysis was limited by the lack of available data beyond those provided by the ERSPC study,[4] the Gothenburg study,[15] and the early phases of the Norrkoping study.[26 27] As none of the studies used age as a stratification factor for randomisation, these subgroup analyses should be interpreted with caution.[40] Our analysis failed to show a clinically important benefit of screening on death from prostate cancer in any of the age groups. The US Preventive Services Task Force advises against routine prostate cancer screening in patients of age 75 and older to avoid unnecessary testing and overtreatment.[41] Indeed, according to the ERSPC[4] study, prostate cancer screening in men aged 70-74 was associated with a relative risk of death from prostate cancer of 1.26 with a wide 95% confidence interval from 0.80 to 1.99. In terms of relative risk, this could mean either a 20% reduced or about 100% increased risk, therefore caution is justified when recommending routine screening in this population. Our systematic review indicates that similar restraints might be indicated for routine screening in all age groups.

Comparison with other reviews

Our review differs from the previously published Cochrane review[12] in that it includes a total of 387 286 participants and provides a relatively precise estimate of the impact of prostate cancer screening on overall mortality with modest quality of evidence. Based on this analysis, it seems unlikely that future large trials in similar populations of participants with digital rectal examination and testing for prostate specific antigen in a screening setting will yield divergent results. In contrast, the confidence interval surrounding the pooled point estimate for death from prostate cancer includes both a 29% relative reduction as well as a 9% relative increase in risk of death secondary to prostate cancer. If we assume an average event rate of death from prostate cancer of 0.78%, as observed in the control arm of the Gothenburg study,[15] screening of 10 000 men would be expected to result in about 22 fewer to seven more deaths related to prostate cancer.

Although more extended follow-up of existing trials (and potential future trials if funded and executed) could help to further characterise the actual effect of screening for prostate cancer on disease specific mortality, our findings suggest that the expected impact in absolute terms would be modest at best.

Study implications

Our study highlights the complexities of the controversy over prostate cancer screening, in particular those of overdiagnosis and the poorly quantifiable downstream harms of overtreatment and impact on quality of life that none of the existing trials has adequately addressed. Accurate estimates of rates of overdiagnosis are challenging, requiring studies with sufficiently long term follow-up data, a criterion not met by any of the available trials on prostate cancer screening, as well as modelling the natural course of prostate cancer, the impact of early diagnosis, and the role of competing mortality.[36] Our systematic review based on all the available evidence failed to show a significant impact of screening on all cause mortality or disease specific mortality. These findings suggest that the rate of overdiagnosis corresponds to the rate of diagnosis of prostate cancer in the screening group compared with the control group.

This study further identifies the challenges faced by future clinical trials. These include the choice of the appropriate screening threshold and interval, the issue of contamination between the screening and control arm, and compliance with recommendations for biopsy. Each of these factors differed considerably between the two major ERSPC and PLCO studies and probably contributed to their different outcomes.[42] It is also noteworthy that these two trials published their results for different reasons. Whereas the ERSPC was published after its third interim analysis showed a significant benefit in favour of screening, the PLCO study was stopped over concerns of potential harm.[42]

Several related studies are ongoing and expected to provide further evidence on the benefits and harms of screening as well as the effect of subsequent treatment choices in patients with positive results. The prostate testing for cancer and treatment (ProtecT) trial in the United Kingdom[43 44] and its extension, the comparison arm for ProtecT (CAP) trial,[45] are ongoing but not expected to report their final results until 2013 and 2015, respectively. This cluster randomised trial has allocated practices with about 460 000 men aged 50-69 to either usual care or population based screening with prostate specific antigen (biopsy if prostate specific antigen ≥3 ng/ml) followed by randomisation of participants diagnosed with prostate cancer (about 1500 expected) to either radical surgery, conformal radiotherapy, or active surveillance. In the US, from 1994 to 2002 the prostate cancer intervention versus observation trial (PIVOT) randomised 731 men from an ethnically diverse background to either radical prostatectomy or active surveillance.[46] Reporting of the final results is expected within the next one to two

years. Lastly, the surveillance therapy against radical treatment (START) trial is an ongoing study based in Canada that is planning to randomise 2130 men with low risk localised prostate cancer to active surveillance versus early interventions with curative intent.[47] These trials, along with the complete follow-up and full reporting of the PLCO trial,[14] ERSPC,[4] French ERSPC,[13] and Gothenburg trial,[15] will contribute additional valuable information to our knowledge about the benefits and harms of screening for prostate cancer.

In summary, existing evidence from randomised controlled trials does not support the routine use of screening for prostate cancer, though screening probably aids in earlier diagnosis and helps to detect prostate cancer at an earlier stage. This early detection, which has not been shown to have a significant impact on mortality, comes at the price of additional testing, the risk of overtreatment and downstream adverse effects, and impaired quality of life that currently cannot be precisely quantified. Patients need to be informed about the existing uncertainties; individual patients' values and preferences are key factors in deciding whether to offer screening.[4 14 43 46] At the individual level, it is conceivable that some patients would value early detection, while others might want to avoid the risk of overdiagnosis. Until further evidence accumulates, this systematic review should serve as a basis for the development of evidence based clinical practice guidelines by relevant stakeholder organisations and prompt an update of such guidelines that continue to actively promote routine prostate cancer screening[48] even in the absence of reliable evidence, as reflected by our study findings.

> **WHAT IS ALREADY KNOWN ON THIS TOPIC**
>
> Screening for prostate cancer leads to increased diagnosis and is therefore recommended by clinical practice guidelines
>
> It is unclear whether screening improves overall and disease specific mortality, the most critical outcomes for patients, and whether the overall benefits of screening outweigh the potential harms and costs of overdetection and overtreatment
>
> **WHAT THIS STUDY ADDS**
>
> Screening for prostate cancer does not have a significant impact on either overall mortality or death from prostate cancer
>
> Screening helps to diagnose prostate cancer at an earlier stage but at the risk of overtreatment and downstream adverse effects that currently cannot be precisely quantified

1 Parkin D, Bray F, Devesa S. Cancer burden in the year 2000. The global picture. *Eur J Cancer* 2001;37:S4-66.

2 Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010:10.3322/caac.20073.

3 Lin K, Lipsitz R, Miller T, Janakiraman S. Benefits and harms of prostate-specific antigen screening for prostate cancer: an evidence update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2008;149:192-9.

4 Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320-8.

5 Boyle P, Brawley OW. Prostate cancer: current evidence weighs against population screening. *CA Cancer J Clin* 2009;59:220-4.

6 Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. *JAMA* 2009;302:1685-92.

7 Barry MJ. Screening for prostate cancer—the controversy that refuses to die. *N Engl J Med* 2009;360:1351-4.

8 Wolf AMD, Wender RC, Etzioni RB, Thompson IM, D'Amico AV, Volk RJ, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. *CA Cancer J Clin* 2010;60:70-98.

9 Brawley OW, Gansler T. Introducing the 2010 American Cancer Society prostate cancer screening guideline. *CA Cancer J Clin* 2010;60:68-9.

10 American Urological Association. AUA responds to American Cancer Society guideline for the early detection of prostate cancer 2010. American Urological Association. March 3, 2010. www.auanet.org/content/press/press_releases/article.cfm?articleNo=178.

11 Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann Intern Med* 2009;151:203-5.

12 Ilic D, O'Connor D, Green S, Wilt T. Screening for prostate cancer. *Cochrane Database Syst Rev* 2006;3:CD004720.

13 Jegu J, Tretarre B, Grosclaude P, Rebillard X, Bataille V, Malavaud B, et al. Results and participation factors to the European randomized study of screening for prostate cancer (ERSPC) with prostate specific antigen: French departments of Tarn Herault. *Prog Urol* 2009;19:487-98.

14 Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009;360:1310-9.

15 Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, et al. Mortality results from the Göteborg randomised population-based prostate-cancer screening trial. *Lancet Oncol* 2010;11:725-32.

16 Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.

17 American Joint Committee on Cancer. Cancer staging manual. 7th ed. Springer, 2010.

18 Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. Cochrane Collaboration and John Wiley, 2008.

19 Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, et al. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008;336:995-8.

20 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34.

21 Review manager (Revman) [program]. Nordic Cochrane Center, Cochrane Collaboration, 2008.

22 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.

23 Hoffman R. Randomized trial results did not resolve controversies surrounding prostate cancer screening. *Curr Opin Urol* 2010;20:189-93.

24 Brawley OW, Ankerst DP, Thompson IM. Screening for prostate cancer. *CA Cancer J Clin* 2009;59:264-73.

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

26 Sandblom G, Varenhorst E, Löfman O, Rosell J, Carlsson P. Clinical consequences of screening for prostate cancer: 15 years follow-up of a randomised controlled trial in Sweden. *Eur Urol* 2004;46:717-24.

27 Varenhorst E, Carlsson P, Capik E, Lofman O, Pudersen K. Repeated screening for carcinoma of the prostate by digital rectal examination in a randomly selected population. *Acta Oncol* 1992;31:815-21.

28 Labrie F, Candas B, Cusan L, Gomez JL, Belanger A, Brousseau G, et al. Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. *Prostate* 2004;59:311-8.

29 Labrie F, Candas B, Dupont A, Cusan L, Gomez JL, Suburu RE, et al. Screening decreases prostate cancer death: first analysis of the 1988 Quebec prospective randomized controlled trial. *Prostate* 1999;38:83-91.

30 Labrie F, Cusan L, Gomez J, Levesque J, Candas B. Screening and treatment of localized prostate cancer decreases mortality: first analysis of the first prospective and randomized study on prostate cancer screening. *Aging Male* 1999;2:33-43.

31 Makinen T, Auvinen A, Hakama M, Stenman UH, Tammela TL. Acceptability and complications of prostate biopsy in population-based PSA screening versus routine clinical practice: a prospective controlled study. *Urology* 2002;60:846-50.

32 Raaijmakers R, Kirkels WK, Roobol MJ, Wildhagen MF, Schroder FH. Complication rates and risk factors of 5802 transrectal ultrasound-guided sextant biopsies of the prostate within a population-based screening program. *Urology* 2002;60:826-30.

33 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.

34 Albertsen PC, Fryback DG, Storer BE, Kolon TF, Fine J. Long-term survival among men with conservatively treated localized prostate cancer. *JAMA* 1995;274:626-31.

35 Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. *JAMA* 2005;293:2095-101.

36 Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:1-9.

37 Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst* 2002;94:981-90.

38 Draisma G, Boer R, Otto SJ, van der Cruijsen IW, Damhuis RAM, Schroder FH, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European randomized study of screening for prostate cancer. *J Natl Cancer Inst* 2003;95:868-78.

39 Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Gulati R, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst* 2009;101:374-83.

40 Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010;340:c117.

41 US Preventive Services Task Force. Screening for prostate cancer: US Preventive Services Task Force recommendation statement. *Ann Intern Med* 2008;149:185-91.

42 Schroder FH, Roobol MJ. ERSPC and PLCO prostate cancer screening studies: What are the differences? *Eur Urol* 2010;58:46-52.

43 Donovan J, Hamdy F, Neal D, et al. Prostate testing for cancer and treatment (ProtecT) feasibility study. *Health Technol Assess* 2003;7:1-81.

44 The ProtecT trial. Evaluating the effectiveness of treatment for clinically localised prostate cancer 2007. ISRCTN Register. 11 June 2010. www.controlled-trials.com/ISRCTN20141297.

45 Evaluating population-based screening for localised prostate cancer in the United Kingdom: an extension to the ProtecT treatment trial 2006. *ISRCTN Register* 17 March 2010. www.controlled-trials.com/ISRCTN92187251.

46 Wilt TJ, Brawer MK, Barry MJ, Jones TM, Kwon Y, Gingrich JR, et al. The prostate cancer intervention versus observation trial: VA/NCI/AHRQ Cooperative Studies Program #407 (PIVOT): design and baseline results of a randomized controlled trial comparing radical prostatectomy to watchful waiting for men with clinically localized prostate cancer. *Contemp Clin Trials* 2009;30:81-7.

47 Wilt TJ. SPCG-4: a needed START to PIVOTal data to promote and ProtecT evidence-based prostate cancer care. *J Natl Cancer Inst* 2008;100:1123-5.

48 Kawachi MH, Bahnson RR, Barry M, Busby JE, Carroll PR, Carter HB, et al. National Comprehensive Cancer Network clinical practice guidelines in oncology: Prostate cancer early detection. *J Natl Compr Canc Netw* 2010;8:240-62.