Benefits and Harms of Prostate-Specific Antigen Screening for Prostate Cancer: An Evidence Update for the U.S. Preventive Services Task Force

Kenneth Lin, MD; Robert Lipsitz, MD, MPH; Therese Miller, DrPH; and Supriya Janakiraman, MD, MPH

Background: Prostate cancer is the most common nonskin cancer in men in the United States, and prostate cancer screening has increased in recent years. In 2002, the U.S. Preventive Services Task Force concluded that evidence was insufficient to recommend for or against screening for prostate cancer with prostate-specific antigen (PSA) testing.

Purpose: To examine new evidence on benefits and harms of screening asymptomatic men for prostate cancer with PSA.

Data Sources: English-language articles identified in PubMed and the Cochrane Library (search dates, January 2002 to July 2007), reference lists of retrieved articles, and expert suggestions.

Study Selection: Randomized, controlled trials and meta-analyses of PSA screening and cross-sectional and cohort studies of screening harms and of the natural history of screening-detected cancer were selected to answer the following questions: Does screening for prostate cancer with PSA, as a single-threshold test or as a function of multiple tests over time, decrease morbidity or mortality? What are the magnitude and nature of harms associated with prostate cancer screening, other than overtreatment? What is the natural history of PSA-detected, nonpalpable, localized prostate cancer?

Data Extraction: Studies were reviewed, abstracted, and rated for quality by using predefined U.S. Preventive Services Task Force criteria.

Data Synthesis: No good-quality randomized, controlled trials of screening for prostate cancer have been completed. In 1 cross-sectional and 2 prospective cohort studies of fair to good quality, false-positive PSA screening results caused psychological adverse effects for up to 1 year after the test. The natural history of PSA-detected prostate cancer is poorly understood.

Limitations: Few eligible studies were identified. Long-term adverse effects of false-positive PSA screening test results are unknown.

Conclusion: Prostate-specific antigen screening is associated with psychological harms, and its potential benefits remain uncertain.

Ann Intern Med. 2008;149:192-199. www.annals.org
For author affiliations, see end of text.

Prostate cancer is the most common nonskin cancer in U.S. men. An estimated 218 890 men received a new diagnosis of prostate cancer in 2007, and 1 in 6 men will receive a diagnosis in their lifetime (1). The American Cancer Society estimates that 27 350 men died of prostate cancer in 2006 (2). After peaking in 1991 (29.4 deaths per 100 000 men), the prostate cancer mortality rate has gradually decreased. Although this positive trend may be related to increased screening for prostate cancer, other factors, including new treatment approaches, could also account for some or all of the observed decline in mortality (3).

The serum prostate-specific antigen (PSA) test was approved by the U.S. Food and Drug Administration in 1986, and its use for prostate cancer screening has increased substantially since the mid-1990s (4). However, PSA testing is not specific to prostate cancer; common conditions, such as benign prostatic hyperplasia and prostatitis, also increase PSA levels. Approximately 1.5 million U.S. men age 40 to 69 years have a PSA level greater than 4.0 μg/L, a widely used cutoff value for a positive screening result (5). Refinements designed to improve the PSA test's sensitivity and specificity for prostate cancer include determination of PSA density, PSA velocity, PSA doubling time, and percentage of free PSA (6–9).

Potential harms from PSA screening include additional medical visits, adverse effects of prostate biopsies, anxiety, and overdiagnosis (the identification of prostate cancer that would never have caused symptoms in the patient's lifetime, leading to unnecessary treatment and associated adverse effects). Much uncertainty surrounds which cases of prostate cancer require treatment and whether earlier detection leads to improvements in duration or quality of life. Two recent systematic reviews of the comparative effectiveness and harms of therapies for localized prostate cancer concluded that no single therapy is superior to all others in all situations (10, 11).

In 2002, the U.S. Preventive Services Task Force (USPSTF) found insufficient evidence to recommend for or against routine screening for prostate cancer. The USPSTF found good evidence that PSA screening can de-

tect early-stage prostate cancer but found mixed and inconclusive evidence that screening and early detection improve health outcomes. Consequently, the USPSTF was unable to determine the balance between benefits and harms of periodic screening for prostate cancer.

The analytic framework that guided the previous USPSTF evidence review (**Figure**) (12) included 8 key questions about benefits and harms of prostate cancer screening and treatment. This evidence update focuses on critical gaps in the evidence that the Task Force identified in the previous review: the lack of good-quality studies linking screening to improved health outcomes; limited information about harms of screening; and a paucity of knowledge about the natural history of PSA-detected, nonpalpable, localized prostate cancer (the most common type of prostate cancer detected today). These evidence gaps produced 3 new key questions for this update:

1. Does screening for prostate cancer with PSA, as a single-threshold test or as a function of multiple tests over time, decrease morbidity or mortality?

2. What are the magnitude and nature of harms associated with prostate cancer screening, other than overtreatment?

3. What is the natural history of PSA-detected, nonpalpable, localized prostate cancer?

METHODS

After consultation with USPSTF liaisons and content experts, we chose a broad definition of *PSA screening* that included evolving prognostic measures, such as PSA velocity and doubling time. However, a comparison of the performance characteristics of such measures with traditional single-threshold PSA testing is outside the scope of this review.

Data Sources

For evidence on health outcomes associated with PSA screening, we searched PubMed for English-language articles indexed between 1 January 2002 and 12 July 2007 by using combinations of the Medical Subject Heading (MeSH) terms and keywords *prostate neoplasms*, *screening*, *prostate-specific antigen*, *early diagnosis*, *PSA velocity*, *PSA doubling time*, and *prostate specific antigen doubling*.

For evidence on the harms of screening for prostate cancer, we searched PubMed for English-language articles indexed between 1 January 2002 and 12 July 2007 by using combinations of the MeSH terms and keywords *prostate neoplasms*; *screening*; *false positive reactions*; *adverse effects*; *mass screening/adverse effects*; *mass screening/psychology*; *anxiety*; *quality of life*; and *health knowledge, attitudes, practice*.

For evidence on the natural history of PSA-detected, nonpalpable, localized prostate cancer, we searched PubMed for English-language articles indexed between 1 January 2002 and 23 August 2007 by using combinations of the MeSH terms and keywords *prostatic neoplasms*, *nat-*

Figure. Analytic framework for screening for prostate cancer.
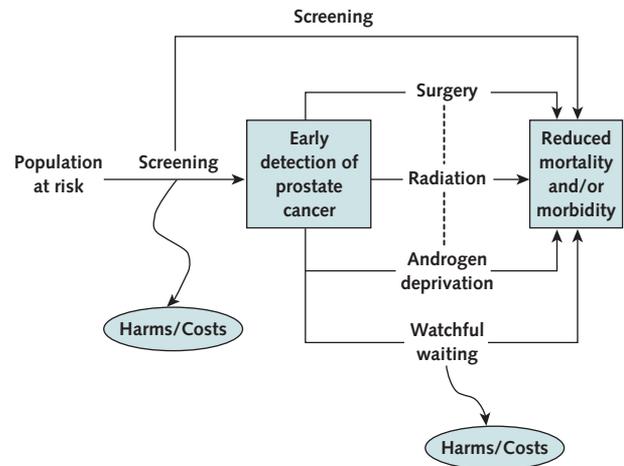
ural history, *epidemiology*, *disease progression*, *survival analysis*, *watchful waiting*, *active surveillance*, *population surveillance*, *expectant management*, and *conservative management*.

We identified additional articles through a search of the Cochrane Library, recommendations of experts, and a hand search of reference lists from major review articles and studies.

Study Selection

Two reviewers independently reviewed the title lists, abstracts, and full articles by using predetermined inclusion and exclusion criteria. Articles selected by at least 1 reviewer advanced to the next stage of review.

For key question 1, eligible studies were randomized, controlled trials (RCTs), meta-analyses, and systematic reviews that compared screening with no screening (or usual care) in general primary care populations and reported morbidity or mortality outcomes. Although the 2002 USPSTF review (12) considered case–control studies and ecological data related to this key question, we excluded these study types from this part of the evidence update to avoid potential sources of confounding that are inherent in nonrandomized studies.

For key question 2, eligible studies were randomized or nonrandomized comparative studies that reported quantitative health or quality-of-life outcomes related to a false-positive screening result. We excluded studies that reported only harms resulting from prostate cancer treatment.

For key question 3, eligible studies were RCTs and cohort studies that reported health outcomes of patients with stage T1c (nonpalpable, localized, PSA-detected) prostate cancer who did not receive active treatment (including patients assigned to watchful waiting or active surveillance protocols). To ensure that we retrieved the most applicable information on natural history, we excluded studies that predominantly involved patients with non–

PSA-detected cancer (defined as comprising ≥80% of the study population), were too small to draw reliable conclusions about health outcomes (defined as <50 patients in the watchful waiting or surveillance group), or did not provide separate data on patients with stage T1c prostate cancer.

Data Extraction and Quality Assessment

For all citations that met the initial eligibility criteria, 2 reviewers reviewed the full articles and independently rated their quality by using previously published USPSTF criteria (13). Disagreements between reviewers regarding article inclusion and quality rating were resolved through a consensus process. We assessed the quality of RCTs and cohort studies on the following items: initial assembly and maintenance of comparable groups; absence of important differential loss to follow-up or overall high loss to follow-up; use of equal, valid, and reliable outcome measurements; clear definition of interventions; and appropriateness of outcomes. We evaluated systematic reviews and meta-analyses on the following items: comprehensiveness of sources considered, appropriateness of search strategy, standard appraisal of included studies, validity of conclusions, recency, and relevance. The **Appendix Table** (available at www.annals.org) describes more thoroughly the criteria and definitions for USPSTF quality ratings.

Data Synthesis and Analysis

We synthesized the data qualitatively by key question in tabular and narrative formats. Data from the 2002 USPSTF review (12) relevant to key questions 1 and 2 are included to facilitate an overall assessment of the body of evidence. We did not perform quantitative synthesis because of the paucity and heterogeneity of included studies.

Role of the Funding Source

Results

We identified 390 potentially relevant articles on health outcomes associated with PSA screening, 421 potentially relevant articles on harms of prostate cancer screening, and 91 potentially relevant articles on the natural history of PSA-detected prostate cancer. **Appendix Figures 1**, **2**, and **3** (available at www.annals.org) contain details of the stages of review and reasons for study exclusion. We obtained 68 articles for full-text review; 10 articles met inclusion criteria for this evidence update.

Key Question 1

Does screening for prostate cancer with PSA, as a single-threshold test or as a function of multiple tests over time, decrease morbidity or mortality?

No good- or fair-quality RCTs addressed this question. Two poor-quality RCTs with important flaws in design and analysis (**Table 1**) do not show a mortality benefit from PSA screening independently or in a meta-analysis. We identified no RCTs that measured health outcomes from PSA screening by means other than single-threshold tests.

In 2002, the USPSTF identified 1 poor-quality RCT of prostate cancer screening by Labrie and colleagues (14) that did not show a mortality benefit from screening when data were reanalyzed by using an intention-to-screen analysis. A 2004 publication described 3 additional years of follow-up of this study (15). By the end of 1999, 23.6% of

---

Table 1. Evidence for Key Question 1: Effect of Prostate Cancer Screening on Morbidity and Mortality

Author, Year (Reference)	Participants	Monitoring Protocol	Results	Comments	Quality Rating
Labrie et al., 2004 (15)	Men registered on the Québec City area electoral rolls in 1985; 46 486 men age 45–80 y; 23% of invited men actually screened; 7.3% of control participants received screening.	Annual screening by PSA and DRE, 1988–1999; PSA cut-point >3.0 μg/L; positive screening result led to TRUS and biopsy.	Primary outcome was death from prostate cancer; 10 deaths in 7348 men who received screening; 74 deaths in 14 231 unscreened men; 62% lower rate of prostate cancer mortality in screened group.	No sociodemographic comparison of the 2 groups; no intention-to-screen analysis; no information on death rates from other causes.	Poor
Sandblom et al., 2004 (16)	All male residents of Norrköping, Sweden, 1987–1996; men age 50–69 y; 1494 men (every 6th man) invited for screening; 70%–78% received screening; 7532 control participants received usual care; an unknown number received screening.	DRE only in 1987 and 1990; DRE and PSA test in 1993 and 1996; PSA cut-point >4.0 μg/L; positive screening result led to biopsy; confirmed prostate cancer treated by urologist with "standardized management program."	Total and prostate cancer–specific survival did not differ between invited and noninvited groups.	No sociodemographic comparison of the 2 groups; no information on crossovers from control group; study was inadequately powered to detect differences in outcomes of interest.	Poor

DRE = digital rectal examination; PSA = prostate-specific antigen; TRUS = transrectal ultrasonography.

Table 2. Evidence for Key Question 2: Harms of PSA Screening

Author, Year (Reference)	Study Type	Participants	Measurements	Results	Comments	Quality Rating
Brindle et al., 2006 (19)	Prospective cohort	Men age 50–69 y in general practices in the United Kingdom	Before-and-after abnormal PSA results: HADS and SF-12	HADS or SF-12 scores did not differ.	Did not specifically assess effect of false-positive screening results; complete data on 78% of participants	Fair
Katz et al., 2007 (20)	Cross-sectional	Men at university-affiliated primary care practices in Wisconsin, Iowa, and Indiana; 97% white; mostly college-educated	2 mo after testing: Medical Outcomes Study SF-36, SAI-6, questions about prostate cancer–related worry (using 5-point scale), perception of cancer risk, and sexual function	SF-36 or SAI-6 scores did not differ; men with false-positive screening results had increased prostate cancer–related worry and problems with sexual function.	84% of biopsy group and 73% of control group returned surveys; findings persisted after multivariate regression analysis; limited external validity because of demographic characteristics of participants	Fair
McNaughton-Collins et al., 2004 (21); Fowler et al., 2006 (22)	Prospective cohort	Men age ≥40 y at primary care practices of 3 Boston teaching hospitals; 88% white; mostly college-educated	6 wk, 6 mo, and 1 y after testing: survey developed in focus groups with questions about prostate cancer worry, prostate cancer knowledge, self-perceived cancer risk, and biopsy experiences	Biopsy group reported more thinking and worrying about prostate cancer than control participants; biopsy group was more likely than control group to have had another PSA test or biopsy.	More than 85% of men in both groups returned 1-y surveys; limited external validity because of demographic characteristics of participants	Good

HADS = Hospital Anxiety and Depression Scale; PSA = prostate-specific antigen; SAI-6 = State Anxiety Index, Short-Form; SF-12 = Short Form-12 Health Survey; SF-36 = Short Form-36 Health Survey.

the screening-invited group and 7.3% of the control group had actually received screening. Comparing all men who were screened with all men who were not, the authors calculated a relative risk (RR) for death from prostate cancer of 0.385 (95% CI, 0.207 to 0.714) in those who were screened. However, when they compared screening-invited men with noninvited men, they found no mortality difference between the 2 groups (RR, 1.085 [CI, 0.822 to 1.433]).

Labrie and colleagues (14) did not report information on the adequacy of randomization, the demographic composition of the 2 groups, or the characteristics of participants who crossed over from screening to no screening or vice versa. Moreover, they did not indicate whether the assessment of outcomes was blinded. Finally, the inappropriate analysis comparing screened with unscreened cohorts did not adjust for potential confounders.

A poor-quality quasi-RCT by Sandblom and colleagues (16) compared total mortality and prostate-cancer–specific mortality in 1494 men who received digital rectal examination and PSA screening with those in 7532 control

participants. An intention-to-screen analysis found no statistical difference in total mortality or prostate-cancer–specific mortality between the 2 groups. Sandblom and colleagues did not report information on the comparability of the groups, crossovers from the control group, or how the cause of death was assigned. Also, the study was not adequately powered to detect a statistically significant difference in the outcomes of interest.

A Cochrane meta-analysis (17) combined these 2 studies by using an intention-to-screen analysis and found no difference in prostate cancer mortality between men invited to prostate cancer screening and control groups (RR, 1.01 [CI, 0.80 to 1.29]). The authors assessed both studies as having a high risk for bias because of the methodological problems discussed previously.

Key Question 2

What are the magnitude and nature of harms associated with prostate cancer screening, other than overtreatment?

One cross-sectional and 2 prospective cohort studies of fair-to-good quality reported short- and long-term psycho-

logical harms from prostate cancer screening (**Table 2**). Although abnormal screening results did not affect summary measures of anxiety or health-related quality of life, men with false-positive PSA screening test results were more likely to worry specifically about prostate cancer, have a higher perceived risk for prostate cancer, and report problems with sexual function compared with control participants for up to 1 year after the test. In 1 study, 26% of men with false-positive screening results reported moderate-to-severe pain during the prostate biopsy; men with false-positive results were also more likely to undergo repeated PSA testing and additional biopsies.

In 2002, the USPSTF found little evidence on harms associated with prostate cancer screening. Digital rectal examination and prostate biopsy cause discomfort or pain in most men. However, in the initial screening round of an RCT, health-related quality-of-life measures were not negatively affected by false-positive screening test results (18).

Brindle and colleagues (19) administered standardized assessments of anxiety, depression, and mental health to 7344 men who received PSA testing. Of the 855 men with a PSA level greater than an age-specific or numerical threshold, 770 returned for a biopsy and then took the questionnaires again before receiving their biopsy results. Assessment scores did not change in patients with an elevated PSA level. Because some elevated PSA levels were true positive, this study was not able to specifically assess the psychological effect of a false-positive PSA result. It was not clear whether the measures used were sensitive enough to detect changes in mental health related to anxiety specific to prostate cancer. Finally, this study was limited by 2 potential sources of selection bias: Recruited patients were already enrolled in a randomized trial of PSA screening, and more than 20% of participants with abnormal PSA levels were not reevaluated.

Other studies have used prostate-specific measures of harms in addition to generic mental health or quality-of-life scores. Katz and colleagues (20) did a telephone survey of 2 groups of men approximately 2 months after PSA screening. After adjustment for baseline characteristics, men with false-positive screening results were statistically more likely than control participants to worry about getting prostate cancer, have a higher perceived 5-year risk for prostate cancer, and report at least moderate problems with sexual function. This study was limited by the potential of confounding through other sources of psychological differences between the 2 groups (for example, referral vs. primary care patient population) and a lower survey response rate in the control group.

McNaughton-Collins and colleagues (21) compared 167 men who had an abnormal screening result but a benign biopsy specimen with 233 men who had a normal PSA level (defined as <2.5 μg/L). After 6 weeks, 49% of men in the biopsy group reported thinking about prostate cancer "a lot" or "some of the time," compared with 18% of the control group. In addition, 40% of the biopsy group worried "a lot" or "some of the time" about developing prostate cancer compared with 8% of the control group. A total of 26% of men experienced moderate-to-severe pain from the biopsy. For 25% of men, the most recent benign biopsy was their third biopsy or more. Statistically significant differences between the biopsy and control groups in anxiety related to prostate cancer and perceived prostate cancer risk persisted 6 months and 1 year later (22). After 1 year, more men in the biopsy group than in the control group had at least 1 additional PSA test (73% vs. 42%) and another biopsy (15% vs. 1%).

Key Question 3

What is the natural history of PSA-detected, nonpalpable, localized prostate cancer?

Three fair-quality cohort studies with small-to-medium sample sizes, highly self-selected elderly patients, and high drop-out rates show that some men with PSA-detected, nonpalpable, localized (stage T1c) prostate cancer have good health outcomes up to 10 years after diagnosis (**Table 3**). We did not identify any population-based studies in which patients with stage T1c prostate cancer were followed longitudinally with no intervention in order to determine health outcomes resulting from the natural progression of disease.

The USPSTF did not directly examine this question in 2002. Recent studies have used PSA testing and biopsy-based monitoring protocols to identify groups of men with "favorable-risk" prostate cancer who were candidates for delayed treatment if subsequent testing showed biochemical or histologic evidence of disease progression. Triggers for treatment varied by study, and many men in the monitoring groups eventually opted for treatment without any objective signs of disease progression.

Hardie and colleagues (23) tested the feasibility of a surveillance protocol in 80 men (median age, 70.5 years) with localized prostate cancer (stage T1 to T2) who were referred to a single tertiary care center in the United Kingdom from 1993 to 2002. Delayed treatment was recommended on the basis of serial PSA level testing and life expectancy assessments. After a median of 42 months of follow-up, 64 men remained on surveillance, 11 had received delayed treatment, and 5 had died of causes other than prostate cancer. This study was limited by the self-selected nature of participants (representing only 10% of eligible patients during the study enrollment period) and the absence of a standardized PSA-based threshold (absolute value or rate of increase) for initiating treatment.

Roemeling and colleagues (24) studied 64 men (mean age, 68.4 years) who chose watchful waiting and were part of a larger cohort of 293 men with stage T1c or T2 prostate cancer who met favorable risk criteria. After a mean follow-up of 82.4 months (range, 23.8 to 119.9 months), 37 men were living and untreated, 19 had chosen treatment, and 8 had died of causes other than prostate cancer. The same authors examined health outcomes in 278 men

(median age, 69.8 years) who chose an active surveillance protocol (25). After a median follow-up of 3.4 years (range, 1.2 to 6 years), 170 men remained on surveillance, 26 had died of causes other than prostate cancer, and 82 had chosen treatment. Both studies by Roemeling and colleagues were limited by having highly self-selected patient populations and high dropout rates.

DISCUSSION

We found inconclusive evidence from RCTs about the health benefits of screening for prostate cancer with PSA. Although we excluded nonrandomized studies of PSA screening, several case–control studies have been published since the 2002 USPSTF review. These studies, conducted in a variety of settings and populations, have yielded conflicting results about the relationship between PSA screening and prostate cancer–related morbidity and mortality (26–31).

In 2007, Aus and colleagues (32) reported interim results from the ERSPC (European Randomized Study of Screening for Prostate Cancer), an ongoing trial that is designed to detect a mortality difference in men randomly assigned to biennial PSA screening or usual care. In the

19 945 men in this subsection of the study, the authors observed a 49% reduction in the risk for metastatic prostate cancer in the screening group (24 cases compared with 47 cases in the control group) after 10 years of follow-up (32).

We determined that this study did not meet inclusion criteria but brought it to the attention of the USPSTF. The primary outcome, metastatic prostate cancer, is an uncertain surrogate for mortality because of high initial response rates to androgen deprivation therapy and competing causes of death. Also, the criterion used for testing to ascertain this outcome may have resulted in unequal attention to the 2 groups, thereby biasing the results. In the absence of symptoms, bone scans were obtained only in men with a PSA level greater than 20 μg/L. Because patients in the control group with prostate cancer had a higher mean PSA level (90.4 μg/L) than did patients in the screening group (19.8 μg/L), the reported difference in metastatic disease may have been exaggerated in favor of the screening group.

Although we found some evidence that false-positive PSA test results are associated with adverse psychological effects, we cannot determine from the existing studies the

Table 3. Evidence for Key Question 3: Natural History of PSA-Detected, Localized Prostate Cancer

Author, Year (Reference)	Study Type	Participants	Monitoring Protocol	Results	Comments	Quality Rating
Hardie et al., 2005 (23)	Prospective cohort	80 men referred to tertiary care center in United Kingdom, 1993–2002, with stage T1–T2 disease, PSA level ≤20 μg/L, Gleason score ≤7, and fitness for radical prostatectomy or radiotherapy; median age, 70.5 y	PSA and DRE every 3–6 mo for first 2 y, then every 6 mo after; no routine repeated biopsy; delayed treatment decision based on rate of PSA level increase; no established triggers for treatment; decisions "made according to the judgment of each patient and clinician."	Median follow-up, 42 mo; 64 patients remained on surveillance; 5 died of causes other than prostate cancer; 11 received or would receive radical treatment; no evidence of disease recurrence in 7 patients who received treatment.	Self-selected population; only 10% of patients with localized prostate cancer entered protocol; short follow-up.	Fair
Roemeling et al., 2006 (24)	Prospective cohort	64 men with stage T1c or T2 prostate cancer, in the first screening round of the Rotterdam section of ERSPC, who chose watchful waiting; 50 had stage T1c prostate cancer; Gleason score ≤6, PSA level <15 μg/L, and PSA density <0.2 μg/L; mean age, 68.4 y	PSA and DRE every 3 mo for first year, then every 6 mo; "major reason for deferred treatment was an increasing PSA level" but no established triggers.	Mean follow-up, 82.4 mo; 37 patients remained on watchful waiting; 19 patients chose deferred treatment; no deaths from prostate cancer; 8 patients died of other causes.	Characteristics of patients who chose watchful waiting were not separately described.	Fair
Roemeling et al., 2007 (25)	Retrospective cohort	278 men with T1c or T2 prostate cancer in 1 of the 3 screening rounds of the Rotterdam section of ERSPC; median age, 69.8 y; median PSA level, 3.6 μg/L; PSA level <15 μg/L; Gleason score	Not standardized, but consisted of periodic PSA testing; no uniform indications for treatment.	Median follow-up, 3.4 y; 196 patients remained on surveillance; 89% overall survival; 100% prostate cancer–specific survival.	End points (prostate cancer–specific and total mortality) determined by blinded independent committee; short follow-up.	Fair

DRE = digital rectal exam; ERSPC = European Randomized Study of Screening for Prostate Cancer; PSA = prostate-specific antigen.

precise magnitude of psychological harms of prostate cancer screening. Because the populations studied have almost exclusively consisted of college-educated white men, these results may not be generalizable to men with less formal education, or to ethnic or racial minorities. No studies of the effects of false-positive PSA test results have included many black men, who have a higher risk for diagnosis of and death from prostate cancer. Studies excluded during our review that were performed in black patient samples involved strategies for increasing rates of prostate cancer screening in these patients, with the benefit of such screening being assumed.

Short-term monitoring studies of highly selected older men with PSA-detected, nonpalpable, localized prostate cancer do not suggest that delayed or no treatment leads to poor health outcomes. Larger, longer-term studies are urgently needed. A recent report of a population-based cohort of men with untreated, early-stage prostate cancer found a sharp decline in prostate cancer–specific survival after 15 years of follow-up (33). None of the participants in this study received their diagnosis through screening, and most cancer cases were detected when clinically palpable. However, these results, in addition to those from a retrospective cohort study that did not find an increase in prostate cancer mortality rates over a similar period (34), suggest that decades of follow-up may be required to determine the safety and effectiveness of current monitoring protocols.

Although we did not examine new evidence on the harms of treatment in this focused update, the 2002 USPSTF review found that prostate cancer treatments cause clinically significant harms, including erectile dysfunction and urinary incontinence, in many patients (12). Still, many physicians continue to believe that the benefits of immediately treating PSA-detected prostate cancer outweigh the risks of delayed or no treatment. In this context, a study that was excluded from this review merits mention.

In 2005, Bill-Axelson and colleagues (35) reported the results of a trial of 695 men with localized prostate cancer who were randomly assigned to receive radical prostatectomy or watchful waiting. The study did not meet inclusion criteria because only 5.2% of the population had prostate cancer diagnosed through screening and 77.8% of the treatment group had stage T2 (palpable) cancer. After a median of 8.2 years, 14.4% of men in the control group and 8.6% of men in the treatment group had died of prostate cancer. An analysis of prostate cancer–specific mortality stratified by age and intervention suggested that the men younger than 65 years were much more likely to benefit from radical prostatectomy than men 65 years of age or older. In the latter subgroup, the cumulative incidence of death from prostate cancer after 10 years was comparable in the watchful waiting and prostatectomy groups.

Two large RCTs of PSA screening are currently under way. The ERSPC randomly assigned 190 000 men between ages 50 and 75 years to screening with PSA, digital rectal examination, and transrectal ultrasonography or usual care; the intervention was later changed to PSA screening alone (36). Biopsies were performed in patients with PSA levels greater than 3.0 μg/L, and positive biopsy results led to treatment outlined by a standardized protocol. The prostate component of the U.S. National Cancer Institute's Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial is evaluating the effect of annual screening with PSA and digital rectal examination on prostate cancer–specific mortality in 76 705 men (37). Abnormal results were provided to the patient's primary care physician of record, and further diagnostic work-up and treatment were based on individual patient and physician preferences.

These trials may provide valuable and complementary information about the health outcomes associated with PSA screening in the general primary care population. Even if 1 or both ultimately demonstrates a population-level mortality benefit, however, individual screening decisions will still need to be made by weighing the benefits and harms of prostate cancer screening and treatment summarized in the previous USPSTF review (12) and this focused evidence update.

From the Center for Primary Care, Prevention, and Clinical Partnerships, Agency for Healthcare Research and Quality, Rockville, Maryland.

References

1. **National Cancer Institute Surveillance Epidemiology and End Results Program.** Cancer Stat Fact Sheets—Cancer of the Prostate. 2006. Accessed at http://seer.cancer.gov/statfacts/html/prost.html on 5 June 2008.
2. **Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ.** Cancer statistics, 2007. CA Cancer J Clin. 2007;57:43-66. [PMID: 17237035]
3. **Collin SM, Martin RM, Metcalfe C, Gunnell D, Albertsen PC, Neal D, et al.** Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study. Lancet Oncol. 2008;9:445-52. [PMID: 18424233]
4. **Farwell WR, Linder JA, Jha AK.** Trends in prostate-specific antigen testing from 1995 through 2004. Arch Intern Med. 2007;167:2497-502. [PMID: 18071173]
5. **Welch HG, Schwartz LM, Woloshin S.** Prostate-specific antigen levels in the United States: implications of various definitions for abnormal. J Natl Cancer Inst. 2005;97:1132-7. [PMID: 16077071]
6. **Aus G, Damber JE, Khatami A, Lilja H, Stranne J, Hugosson J.** Individualized screening interval for prostate cancer based on prostate-specific antigen level: results of a prospective, randomized, population-based study. Arch Intern Med. 2005;165:1857-61. [PMID: 16157829]
7. **Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A,**

et al. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. JAMA. 1998;279:1542-7. [PMID: 9605898]

8. Catalona WJ, Richie JP, deKernion JB, Ahmann FR, Ratliff TL, Dalkin BL, et al. Comparison of prostate specific antigen concentration versus prostate specific antigen density in the early detection of prostate cancer: receiver operating characteristic curves. J Urol. 1994;152:2031-6. [PMID: 7525994]

9. Fang J, Metter EJ, Landis P, Carter HB. PSA velocity for assessing prostate cancer risk in men with PSA levels between 2.0 and 4.0 ng/ml. Urology. 2002; 59:889-93; discussion 893-4. [PMID: 12031375]

10. Thompson I, Thrasher JB, Aus G, Burnett AL, Canby-Haglino ED, Cookson MS, et al. AUA Prostate Cancer Clinical Guideline Update Panel. Guideline for the management of clinically localized prostate cancer: 2007 update. J Urol. 2007;177:2106-31. [PMID: 17509297]

11. Wilt TJ, MacDonald R, Rutks I, Shamliyan TA, Taylor BC, Kane RL. Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. Ann Intern Med. 2008;148:435-48. [PMID: 18252677]

12. Harris R, Lohr KN. Screening for prostate cancer: an update of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2002;137:917-29. [PMID: 12458993]

13. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Methods Work Group, Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001;20:21-35. [PMID: 11306229]

14. 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. [PMID: 9973093]

15. Labrie F, Candas B, Cusan L, Gomez JL, Bélanger 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. [PMID: 15042607]

16. 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;46:717-23; discussion 724. [PMID: 15548438]

17. Ilic D, O'Connor D, Green S, Wilt T. Screening for prostate cancer. Cochrane Database Syst Rev. 2006;3:CD004720. [PMID: 16856057]

18. Essink-Bot ML, de Koning HJ, Nijs HG, Kirkels WJ, van der Maas PJ, Schröder FH. Short-term effects of population-based screening for prostate cancer on health-related quality of life. J Natl Cancer Inst. 1998;90:925-31. [PMID: 9637143]

19. Brindle LA, Oliver SE, Dedman D, Donovan JL, Neal DE, Hamdy FC, et al. Measuring the psychosocial impact of population-based prostate-specific antigen testing for prostate cancer in the UK. BJU Int. 2006;98:777-82. [PMID: 16978272]

20. Katz DA, Jarrard DF, McHorney CA, Hillis SL, Wiebe DA, Fryback DG. Health perceptions in patients who undergo screening and workup for prostate cancer. Urology. 2007;69:215-20. [PMID: 17320653]

21. McNaughton-Collins M, Fowler FJ Jr, Caubet JF, Bates DW, Lee JM, Hauser A, et al. Psychological effects of a suspicious prostate cancer screening test followed by a benign biopsy result. Am J Med. 2004;117:719-25. [PMID: 15541320]

22. Fowler FJ Jr, Barry MJ, Walker-Corkery B, Caubet JF, Bates DW, Lee JM, et al. The impact of a suspicious prostate biopsy on patients' psychological, sociobehavioral, and medical care outcomes. J Gen Intern Med. 2006;21:715-21.

[PMID: 16808772]

23. Hardie C, Parker C, Norman A, Eeles R, Horwich A, Huddart R, et al. Early outcomes of active surveillance for localized prostate cancer. BJU Int. 2005; 95:956-60. [PMID: 15839912]

24. Roemeling S, Roobol MJ, Postma R, Gosselaar C, van der Kwast TH, Bangma CH, et al. Management and survival of screen-detected prostate cancer patients who might have been suitable for active surveillance. Eur Urol. 2006;50: 475-82. [PMID: 16713065]

25. Roemeling S, Roobol MJ, de Vries SH, Wolters T, Gosselaar C, van Leenders GJ, et al. Active surveillance for prostate cancers detected in three subsequent rounds of a screening trial: characteristics, PSA doubling times, and outcome. Eur Urol. 2007;51:1244-50; discussion 1251. [PMID: 17161520]

26. Weinmann S, Richert-Boe K, Glass AG, Weiss NS. Prostate cancer screening and mortality: a case-control study (United States). Cancer Causes Control. 2004;15:133-8. [PMID: 15017125]

27. Weinmann S, Richert-Boe KE, Van Den Eeden SK, Enger SM, Rybicki BA, Shapiro JA, et al. Screening by prostate-specific antigen and digital rectal examination in relation to prostate cancer mortality: a case-control study. Epidemiology. 2005;16:367-76. [PMID: 15824554]

28. Kopec JA, Goel V, Bunting PS, Neuman J, Sayre EC, Warde P, et al. Screening with prostate specific antigen and metastatic prostate cancer risk: a population based case-control study. J Urol. 2005;174:495-9; discussion 499. [PMID: 16006879]

29. Concato J, Wells CK, Horwitz RI, Penson D, Fincke G, Berlowitz DR, et al. The effectiveness of screening for prostate cancer: a nested case-control study. Arch Intern Med. 2006;166:38-43. [PMID: 16401808]

30. Agalliu I, Weiss NS, Lin DW, Stanford JL. Prostate cancer mortality in relation to screening by prostate-specific antigen testing and digital rectal examination: a population-based study in middle-aged men. Cancer Causes Control. 2007;18:931-7. [PMID: 17641982]

31. Marcella SW, Rhoads GG, Carson JL, Merlino F, Wilcox H. Prostate-specific antigen screening and mortality from prostate cancer. J Gen Intern Med. 2008;23:248-53. [PMID: 18172740]

32. Aus G, Bergdahl S, Lodding P, Lilja H, Hugosson J. Prostate cancer screening decreases the absolute risk of being diagnosed with advanced prostate cancer—results from a prospective, population-based randomized controlled trial. Eur Urol. 2007;51:659-64. [PMID: 16934392]

33. Johansson JE, Andrén O, Andersson SO, Dickman PW, Holmberg L, Magnuson A, et al. Natural history of early, localized prostate cancer. JAMA. 2004;291:2713-9. [PMID: 15187052]

34. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. JAMA. 2005;293:2095-101. [PMID: 15870412]

35. Bill-Axelson A, Holmberg L, Ruutu M, Häggman M, Andersson SO, Bratell S, et al. Scandinavian Prostate Cancer Group Study No. 4. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med. 2005; 352:1977-84. [PMID: 15888698]

36. Schröder FH, Kranse R, Rietbergen J, Hoedemaeke R, Kirkels W. The European Randomized Study of Screening for Prostate Cancer (ERSPC): an update. Members of the ERSPC, Section Rotterdam. Eur Urol. 1999;35:539-43. [PMID: 10408951]

37. Gohagan JK, Prorok PC, Hayes RB, Kramer BS. Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. Control Clin Trials. 2000;21:251S-272S. [PMID: 11189683]

Annals of Internal Medicine

Current Author Addresses: Drs. Lin, Lipsitz, Miller, and Janakiraman:
540 Gaither Road, Rockville, MD 20850.

Hierarchy of research design
 I: Properly conducted RCT
 II-1: Well-designed controlled trial without randomization
 II-2: Well-designed cohort or case–control analytic study
 II-3: Multiple time series with or without the intervention; dramatic results from uncontrolled experiments
 III: Opinions of respected authorities, based on clinical experience; descriptive studies or case reports; reports of expert committees

Design-specific criteria and quality category definitions
 Systematic reviews
  Criteria
   Comprehensiveness of sources considered/search strategy used
   Standard appraisal of included studies
   Validity of conclusions
   Recency and relevance are especially important for systematic reviews
  Definition of ratings based on above criteria
   Good: Recent, relevant review with comprehensive sources and search strategies; explicit and relevant selection criteria; standard appraisal of included
     studies; and valid conclusions
   Fair: Recent, relevant review that is not clearly biased but lacks comprehensive sources and search strategies
   Poor: Outdated, irrelevant, or biased review without systematic search for studies, explicit selection criteria, or standard appraisal of studies

 Case–control studies
  Criteria
   Accurate ascertainment of cases
   Nonbiased selection of cases/controls, with exclusion criteria applied equally to both
   Response rate
   Diagnostic testing procedures applied equally to each group
   Measurement of exposure accurate and applied equally to each group
   Appropriate attention to potential confounding variables
  Definition of ratings based on above criteria
   Good: Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls;
     response rate ≥80%; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to
     confounding variables
   Fair: Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rates <80% or attention to some but not all
     important confounding variables
   Poor: Major section or diagnostic work-up biases, response rates <50%, or inattention to confounding variables

 RCTs and cohort studies
  Criteria
   Initial assembly of comparable groups
    For RCTs: Adequate randomization, including first concealment and whether potential confounders were distributed equally among groups
    For cohort studies: Consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of
      inception cohorts
   Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination)
   Important differential loss to follow-up or overall high loss to follow-up
   Measurements: equal, reliable, and valid (includes masking of outcome assessment)
   Clear definition of the interventions
   All important outcomes considered
  Definition of ratings based on above criteria
   Good: Evaluates relevant available screening tests; uses a credible reference standard; interprets reference standard independently of screening test;
     reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (>100) of broad-spectrum
     patients
   Fair: Evaluates relevant available screening tests; uses reasonable although not best standard; interprets reference standard independent of screening test;
     moderate sample size (50–100 participants) and a "medium" spectrum of patients
   Poor: Has fatal flaw, such as uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very
     small sample size or very narrow selected spectrum of patients

 Diagnostic accuracy studies
  Criteria
   Screening test relevant, available for primary care, adequately described
   Study uses a credible reference standard, performed regardless of test results
   Reference standard interpreted independently of screening test
   Handles indeterminate result in a reasonable manner
   Spectrum of patients included in study
   Sample size
   Administration of reliable screening test
  Definition of ratings based on above criteria
   Good: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability
     of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (>100) of broad-spectrum patients with and
     without disease
   Fair: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test;
     moderate sample size (50–100 participants) and a "medium" spectrum of patients
   Poor: Has fatal flaw, such as uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very
     small sample size or very narrow selected spectrum of patients

Appendix Figure 1. Key question 1: stages of article review.

Articles identified by initial literature review
(*n* = 390)

→ Articles excluded at title stage
(*n* = 342)

Abstracts reviewed
(*n* = 48)

→ Articles excluded at abstract stage
(*n* = 42)

Complete articles reviewed
(*n* = 6)

→ Articles excluded at complete review stage
(*n* = 3)

Articles meeting inclusion criteria (*n* = 3)
Ilic et al. (2006)
Labrie et al. (2004)
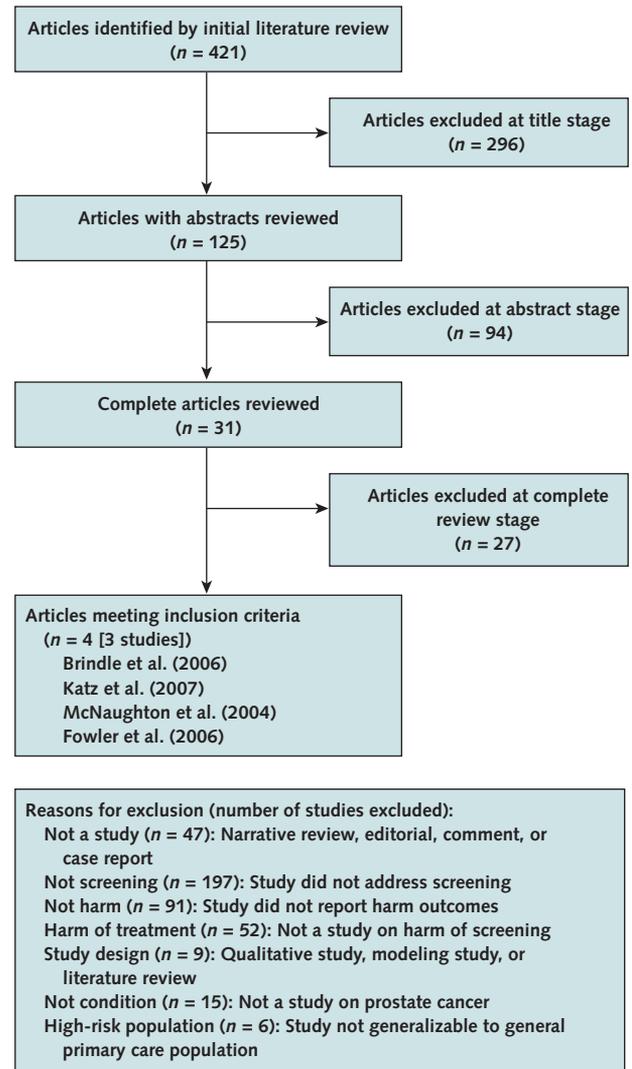Sandblom et al. (2004)

Reasons for exclusion (number of studies excluded):
Not a study (*n* = 6): Narrative review, editorial, comment, or case report
Not screening (*n* = 318): Study did not address screening
Study design (*n* = 1): Not an RCT, meta-analysis, or systematic review
Not condition (*n* = 3): Not a study on prostate cancer
No outcomes (*n* = 59): Contains no information on health outcomes of interest

RCT = randomized, controlled trial.

Appendix Figure 2. Key question 2: stages of article review.

Articles identified by initial literature review
(*n* = 421)

→ Articles excluded at title stage
(*n* = 296)

Articles with abstracts reviewed
(*n* = 125)

→ Articles excluded at abstract stage
(*n* = 94)

Complete articles reviewed
(*n* = 31)

→ Articles excluded at complete review stage
(*n* = 27)

Articles meeting inclusion criteria
(*n* = 4 [3 studies])
Brindle et al. (2006)
Katz et al. (2007)
McNaughton et al. (2004)
Fowler et al. (2006)

Reasons for exclusion (number of studies excluded):
Not a study (*n* = 47): Narrative review, editorial, comment, or case report
Not screening (*n* = 197): Study did not address screening
Not harm (*n* = 91): Study did not report harm outcomes
Harm of treatment (*n* = 52): Not a study on harm of screening
Study design (*n* = 9): Qualitative study, modeling study, or literature review
Not condition (*n* = 15): Not a study on prostate cancer
High-risk population (*n* = 6): Study not generalizable to general primary care population

Appendix Figure 3. **Key question 3: stages of article review.**

Articles with abstracts identified by
literature review
(*n* = 91)

Articles excluded at abstract stage
(*n* = 60)

Complete articles reviewed
(*n* = 31)

Articles excluded at complete
review stage
(*n* = 28)

Articles that met inclusion criteria and were
at least fair quality (*n* = 3)

Reasons for exclusion (number of studies excluded):
Not a study (*n* = 25): Narrative review, editorial, comment, or
    case report
No outcomes (*n* = 27): Study did not report health outcomes
Active treatment (*n* = 10): All patients received some form of
    active treatment
Study design (*n* = 7): An excluded study design, or a design not
    capable of answering this key question (for example, does not
    report T1c cancer outcomes separately)
Not condition (*n* = 16): Not a study on prostate cancer
Duplicate study (*n* = 1): No new information additional to that in
    a previous publication
Quality (*n* = 2): Rated as poor quality by both reviewers