Screening for Prostate Cancer: An Update of the Evidence

Russell Harris, MD, MPH; Kathleen N. Lohr, PhD

**Epidemiology**

The American Cancer Society estimates that 189,000 men will be diagnosed with prostate cancer in 2002 and that 30,200 men will die of the disease. Many more men receive a diagnosis of prostate cancer than die of it (lifetime risk, about 1 in 6 vs. about 1 in 29). Among types of cancer, only lung cancer kills more men each year. The cause of prostate cancer is unknown, and the best-documented risk factors (age, ethnicity, and family history) are not modifiable. The burden of prostate cancer falls disproportionately on men who are older or black. The median age at diagnosis is approximately 71 years, and the median age at death is 78 years. More than 75% of all cases of prostate cancer are diagnosed in men older than 65 years of age, and 90% of deaths occur in this age group. Incidence is approximately 60% higher and mortality rate is twofold higher in black men than in white men. Asian-American men and Hispanic men have incidence rates lower than non-Hispanic white persons.

Although approaches to primary prevention of prostate cancer are being tested, to date none are known to be effective. The most common strategy for reducing the burden of prostate cancer is screening, but screening remains controversial. Many studies on this topic have been published since 1996, when the U.S. Preventive Services Task Force (USPSTF) last examined prostate screening. To assist the USPSTF in updating its recommendation, the RTI-University of North Carolina Evidence-based Practice Center performed a systematic review of the evidence on screening for prostate cancer.

**Methods**

Using USPSTF methods, we developed an analytic framework and eight key questions to guide our literature search. Because we found no direct evidence connecting screening and reduced mortality, we searched for indirect evidence concerning the yield of screening, the efficacy and harms of various forms of treatment for early prostate cancer, and the costs and cost-effectiveness of screening. We developed eligibility criteria for selecting the evidence relevant to answer the key questions (Table 1). We examined the critical literature from the 1996
USPSTF review and used search terms consistent with the eligibility criteria to search the MEDLINE database and Cochrane Library for English language reviews and relevant studies published between January 1, 1994, and September 15, 2002.

The first author and at least one trained assistant reviewed abstracts and articles to find those that met the eligibility criteria. For these included studies, the two reviewers abstracted relevant information using standardized abstraction forms. We graded the quality of all included articles according to USPSTF criteria. The authors worked closely with two members of the USPSTF throughout the review and periodically presented reports to the full USPSTF. We distributed a draft of the systematic evidence review to experts in the field and relevant professional organizations and federal agencies for broad-based external peer review, and made revisions based on the feedback. We then revised the full systematic evidence review into this manuscript.

A more complete account of the methods of this review can be found in the Appendix. The

<table>
<thead>
<tr>
<th>Key question</th>
<th>Inclusion criteria</th>
<th>Number of articles meeting criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Published 1/1/94–9/15/02 English Language MEDLINE, Cochrane Human subjects</td>
<td></td>
</tr>
<tr>
<td>1. Efficacy of screening (direct evidence)</td>
<td>RCT; case-control study; or ecologic evidence directly connecting screening with health outcomes</td>
<td>RCT: 1 Case-control: 2 Ecologic: 15</td>
</tr>
<tr>
<td>2. Yield of screening</td>
<td>• Unselected population without prostate cancer • Screening test offered to all • Work-up offered to all positive screens • Screening test compared to a valid reference standard</td>
<td>35</td>
</tr>
<tr>
<td>3. Efficacy of radical prostatectomy</td>
<td>For Key Questions 3-6: • RCT; clinically localized disease • Follow-up ≥ 2 years</td>
<td>1</td>
</tr>
<tr>
<td>4. Efficacy of radiation therapy</td>
<td>• ≥ 75% of patients followed • Health outcomes</td>
<td>0</td>
</tr>
<tr>
<td>5. Efficacy of androgen deprivation therapy</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>6. Efficacy of watchful waiting</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>7. Harms of treatment</td>
<td>• Patient self-report • Use of valid measurement instrument • Follow-up from pre-treatment to at least 12 months post-treatment, or • Comparison with similar untreated control group at least 12 months post-treatment</td>
<td>32</td>
</tr>
<tr>
<td>8. Costs and cost-effectiveness of treatment</td>
<td>• Valid assessment of costs of screening and treatment • Assess direct and indirect costs • Cost-effectiveness, cost-benefit, cost-utility • Modeling studies</td>
<td>2</td>
</tr>
</tbody>
</table>
complete Systematic Evidence Review is available on the Web site of the Agency for Healthcare Research and Quality (www.ahrq.gov). This evidence report was funded through a contract to the Research Triangle Institute-University of North Carolina Evidence-based Practice Center from the Agency for Healthcare Research and Quality (AHRQ). Staff of the funding agency and members of the USPSTF contributed to the study design, reviewed draft and final manuscripts, and made editing suggestions.

Results

Direct Evidence that Screening Reduces Mortality

Randomized Controlled Trials (RCTs)

Labrie et al. completed the first randomized controlled trial of prostate cancer screening with more than 46,000 men. At the end of 8 years of follow-up, approximately 23% of the invited group and 6.5% of the not-invited group had been screened with prostate-specific antigen (PSA) and digital rectal examination (DRE). Prostate cancer death rates did not differ between groups (4.6 vs. 4.8 deaths per 1,000 people, respectively).

Two other RCTs of prostate cancer screening, both initiated in 1994, are ongoing: the U.S. National Cancer Institute “Prostate, Lung, Colorectal, and Ovary” Trial and the European Randomized Study of Screening for Prostate Cancer. Neither study will have data on mortality for several more years.

Case-Control Studies

Three well-conducted, nested case-control studies (two since 1994) examined the relationship between chart review documentation of DRE and advanced prostate cancer or death from prostate cancer. Two studies found no relationship. The third study found that men who died of prostate cancer had fewer DREs in the years before diagnosis (odds ratio indicating a protective effect of DRE, [OR] = 0.51; 95% confidence interval [CI], 0.31-0.84).

Why results from these otherwise similar studies differ is not clear. The 3 studies depended on large databases and on individual medical records. They defined cases slightly differently and used different approaches to differentiate screening DRE from diagnostic DRE. Because such studies are complex in design, we were not able to determine whether one method was more accurate than another. All three studies were small, and all were consistent with a reduction in prostate cancer mortality of up to 50% with DRE.

We found no case-control studies of PSA screening. This can be explained, at least in part, by the fact that insufficient time has elapsed since the introduction of PSA as a screening test in the late 1980s. Such studies are under way.

Ecologic Studies

Around 1987, PSA screening began to increase rapidly in the United States. Important trends in prostate cancer incidence and mortality also occurred at that time. Although incidence rates had been slowly increasing for some years before 1987, data from the U.S. Surveillance, Epidemiology, and End Results (SEER) program showed a dramatic increase in age-adjusted prostate cancer incidence-20% per year from 1989 to 1992. The rates then decreased at 10.8% per year, stabilizing after 1994. Most of the increase in incidence was seen in localized or regional disease. Incidence of distant-stage disease at diagnosis showed little initial increase and then began to decline; annual decline for white men was 17.9% after 1991.

Disease-specific mortality rates paralleled trends in prostate cancer incidence. In the late 1980s, the average annual percentage increase rose from 0.7% to 3.1% for white men and from 1.6% to 3.2% for black men. In 1991, prostate cancer mortality rates for white men began to decline (21.6% decrease from 1991 to 1999); in 1993, rates for black men followed suit (16.0% decrease from 1993 to 1999). Mortality rates declined in all age groups at about the same time. Analyses of trends in prostate cancer incidence and mortality in Olmsted County, Minnesota, and in Canada have shown similar results.

Ecologic evidence is difficult to interpret. Although screening probably explains trends in incidence of prostate cancer, trends in mortality are more difficult to understand. Some aspects of the
trends (eg, a decline in distant-stage disease) are consistent with screening, but other aspects (eg, the short time between increased screening and decreased mortality) are not as consistent with our current view of the natural history of prostate cancer. The argument that the decline in mortality can be attributed to PSA screening would be stronger if it could be shown that the decline was largest in areas with more screening. To date, data on this issue are conflicting.

Other possible explanations for decreased mortality include “attribution bias” and improved treatment. Attribution bias suggests that some deaths among men with prostate cancer are mistakenly attributed to the prostate cancer. If the percentage of deaths so attributed is stable, then the prostate cancer mortality rate would be expected to increase and decrease in close approximation with the incidence of prostate cancer in the population.

Changes in prostate cancer treatment during the late 1980s and early 1990s included higher rates of radical prostatectomy, development of luteinizing hormone-releasing hormone (LHRH) agonists (allowing improved androgen deprivation therapy without castration), and refinements in radiation therapy. Such changes are a potential explanation for the reduction in prostate cancer mortality. A recent study by Bartsch et al., for example, documented a greater reduction in prostate cancer mortality in the Austrian state of Tyrol, which had instituted a free PSA screening program, compared with the rest of Austria. This finding could be a consequence of the screening program, changes in treatment that accompanied the screening program, misattribution of the cause of death, or some combination of the three.

Accuracy of Screening

Three problems complicate any attempt to determine the accuracy of screening tests for prostate cancer. First, research has yet to clarify which tumors screening should target. Second, the reference standard (prostate biopsy) for diagnosing prostate cancer after a positive screening test is imperfect. Third, few studies perform biopsy on men with negative results on screening tests.

Prostate cancer is a heterogeneous tumor. Different cases of prostate cancer have widely varying growth rates and potential for causing death. Ideally, prostate cancer screening would target only tumors that would cause clinically important disease. Currently available prognostic markers can distinguish a small number of men with excellent prognosis for long-term survival and a small number of men with poor prognosis for long-term survival. However, they cannot help us correctly categorize the prognosis of those in the middle category, which includes most men with prostate cancer. Since research has not yet clearly defined the characteristics of clinically important prostate cancer, we do not know what the specific target of screening should be.

The usual reference standard test used in prostate cancer screening studies, transrectal needle biopsy of the prostate, is imperfect for two reasons. First, it misses some cases of cancer: 10% to 30% of men who have negative results on an initial series of biopsies have cancer on repeated biopsy series. Thus, some men categorized as not having cancer actually have it, falsely lowering the test’s measured sensitivity.

Second, in clinical practice and research, a “biopsy” is actually four to six (or more) biopsies. Many biopsy specimens are obtained, most from normal-appearing areas of the prostate. An analysis of this practice concluded that up to 25% of apparently PSA-detected tumors and more than 25% of apparently DRE-detected tumors were likely in fact to have been detected by serendipity, that is, an incidental finding from a blind biopsy. Thus, some men who are categorized as having cancer detected by screening actually have serendipity-detected cancer. This error falsely increases sensitivity.

In addition to problems of the accuracy of the reference standard, few studies perform biopsy on men who have negative results on screening tests. This reduces our ability to determine the number of false-negative screening tests and to calculate sensitivity. Most studies use several noninvasive tests together, measuring the sensitivity of one test against the combined findings of all tests. The extent to which these combined tests actually
detect all important cancers is unknown. This bias probably leads to an overestimate of sensitivity.

**Screening Methods**

**PSA screening.** An analysis from the Physicians' Health Study avoided some of the bias of the problematic reference standard by using longitudinal follow-up instead of biopsy. In this study, which used a PSA cutpoint of 4.0 ng/ml or higher, the sensitivity for detecting cancers appearing within two years after screening was 73.2%. Although the study calculated sensitivity separately for aggressive (ie, extracapsular or higher grade) and nonaggressive (ie, intracapsular and lower grade) cancer (sensitivity, 91% vs. 56%), it is not clear that these categories correspond to clinically important and clinically unimportant tumors. Among men who did not receive a diagnosis of prostate cancer in those two years, 14.6% had an initial PSA of 4.0 ng/ml or greater (corresponding to a specificity of 85.4%).

Other studies have provided similar estimates of sensitivity and specificity for PSA with a cutpoint of 4.0 ng/ml. Specificity for PSA screening is lower among men with larger prostate glands, including the large number of older men with benign prostatic hyperplasia (BPH). One study of four carefully chosen samples found that the likelihood ratios for various PSA levels were much lower among men with BPH than among men without BPH. Thus, the PSA test is not as accurate in detecting cancer among men with BPH as in those without.

Because of the reduced specificity in older men with BPH, some experts have proposed that the PSA cutpoint be adjusted for age, with higher cutpoints for older men and lower cutpoints for younger men. Such a strategy increases sensitivity and lowers specificity in younger men, while the reverse is true in older men. Experts disagree about whether this strategy would lead to improved health outcomes.

Some experts have proposed decreasing the cutpoint defining an abnormal PSA level from 4.0 ng/ml to 3.0 ng/ml (or even 2.6 ng/ml). This approach results in more biopsies and more cancer detected. Because of the uncertainty about the definition of clinical importance, the value of this increased detection is unknown.

In the serum, PSA circulates in two forms: free and complexed with such molecules as alpha-1 antichymotrypsin. Men with prostate cancer tend to have a lower percentage of free PSA than do men without prostate cancer. In research, the percentage of free PSA (or a similar test measuring the level of complexed PSA) has mainly been used to increase the specificity of screening by distinguishing between men with PSA levels of 4.0 ng/ml to 9.9 ng/ml who should undergo biopsy and those who should not. Different studies have suggested different percent free PSA cutpoints; a lower cutpoint avoids more biopsies but also misses more cancer. High cutpoints (eg, 25%) would avoid about 20% of biopsies and the probability of cancer at that cutpoint is about 8%. In practice, it is not clear whether this probability would be low enough for men and their physicians to forgo biopsy.

Men with prostate cancer have a greater increase in their PSA level over time than men without cancer. It is unclear, however, whether examining the annual rate of change in PSA level (PSA velocity) improves health outcomes or reduces unnecessary biopsies. Because of intra-individual variation, PSA velocity is useful only in men who have three or more tests of PSA level over a period of 1 to 3 years.

**DRE screening.** It is more difficult to detect cancer with DRE than with PSA. A meta-analysis examining studies of DRE in unselected samples screened by both PSA and DRE found a sensitivity of 59% (64% for the four best studies). DRE detects cancer in some men with PSA levels below 4.0 ng/ml (positive predictive value about 10%, according to one large study) or even 3.0 ng/ml, but the tumors are usually small and well differentiated. DRE has limited reproducibility.

**Yield of Large Screening Programs**

Using six studies of screening with a single PSA or PSA and DRE among large, previously unscreened samples, we were able to estimate the
yield of a new screening program among men in different age groups who have not previously been screened. Figure 1 gives estimates of positive test results and cases of cancer detected for screening with PSA alone or screening with PSA and DRE among men in their 60s. Men in their 50s have fewer positive test results and cases of cancer detected while men in their 70s have more.

The percentage of participants with a PSA of 4.0 ng/ml or higher ranged from about 4% among men in their 50s to about 27% among men in their 70s. The percentage of men who had a PSA of 4.0 ng/ml or higher or an abnormal results on DRE ranged from 15% among younger men to 40% among older men. Few other screening tests have such a high percentage of positive tests.

In the screening studies, some men with abnormal results on screening tests did not undergo biopsy. If we assume that biopsy is performed on all men with an abnormal result on a screening test and that the

---

**Table 1. Estimated yield of screening with PSA (or PSA and DRE), ages 60–69 (prevalence screen)**

<table>
<thead>
<tr>
<th>PSA &lt; 4.0 (and negative DRE)</th>
<th>PSA ≥ 4.0 (or positive DRE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>840–890 (720–840)</td>
<td>110–160* (160–280)†</td>
</tr>
<tr>
<td></td>
<td>Biopsy</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>33–48† Prostate Cancer (40–70)§</td>
<td>77–112 No Prostate Cancer (120–210)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>23–34‡ Pathologically Organ Confined (28–49)**</td>
<td>10–14 Pathologically Extracapsular (12–21)</td>
</tr>
<tr>
<td></td>
<td>Lives Extended (Unknown)</td>
</tr>
</tbody>
</table>

**Note:** The numbers are estimates of the number of positive tests, cancers detected, and pathologically organ-confined cancers from PSA screening of 1,000 men in their 60s who have not been screened before. The numbers in parentheses are estimates from combined screening with both PSA and DRE. The number of cancers detected would be smaller after repeated annual screening.

* For men in their 50s, about 50; for men in their 70s, about 270.
† For men in their 50s, about 150; for men in their 70s, about 400.
‡ For men in their 50s, about 17; for men in their 70s, about 90.
§ For men in their 50s, about 30; for men in their 70s, about 100.
** For men in their 50s, about 12; for men in their 70s, about 63.
† For men in their 50s, about 21; for men in their 70s, about 70.
rate of cancer detection is the same as for men who undergo study biopsy, we estimate that the percentage of all men screened who would have prostate cancer detected would range from approximately 1.5% (PSA screening alone for men in their 50s)\textsuperscript{51} to 10% (PSA and DRE screening for men in their 70s).\textsuperscript{44}

Figure 1 gives general percentages from all studies. In these six studies, biopsies detected cancer in approximately 30% for men with a PSA of 4.0 ng/ml or higher and 20% to 27% for men with either a PSA of 4.0 ng/ml or higher or an abnormal DRE.\textsuperscript{7,49,80,83} The probability of prostate cancer with a PSA of 2.5 ng/ml to 4.0 ng/ml and a negative DRE is also about 20%.\textsuperscript{51} Although the studies found that 60% to 70% of screen-detected cancer is organ confined,\textsuperscript{7,49,80,82,83} they do not provide information about the number of lives extended by detecting either organ-confined or extracapsular tumors.

**Yield with Different Screening Intervals**

Rates of positive results and cancer detection decrease on screening a year after the initial screening round.\textsuperscript{7,49,80,83-85} In one study, approximately 26% of men with a PSA of 4.0 ng/ml or greater had prostate cancer after the first round of screening and approximately 6.2% had cancer after subsequent rounds.\textsuperscript{7} Other studies have concluded that annual screening confers little gain compared with intervals of at least two years,\textsuperscript{41} especially for the 70% of the population with PSA levels of 2.0 ng/ml or less.\textsuperscript{41-87}

**Effectiveness of Current Treatments for Localized Disease**

**Radical Prostatectomy**

Since 1991, radical prostatectomy (RP) has been the most common treatment for clinically localized prostate cancer. It is the initial treatment for more than one-third of newly patients with new diagnoses, most commonly men 75 years or younger.\textsuperscript{5} The procedure is usually performed with curative intent on men who have a life expectancy of at least 10 years.

One well-conducted RCT compared RP with “watchful waiting” (in which treatment is not given initially but reserved for progressive or symptomatic disease) among men with clinically detected prostate cancer.\textsuperscript{88} About 75% of the men in this study had palpable cancerous tumors; few of which were the size usually detected by PSA screening.\textsuperscript{7,84,89} After eight years of follow-up, 7.1% of the RP group had died of prostate cancer compared with 13.6% of men in the watchful waiting group (relative hazard 0.50; 95% confidence interval [CI] 0.27, 0.91). The absolute difference in prostate cancer mortality was 6.6% (CI 2.1%, 11.1%) (number needed to treat for benefit, 17). The groups did not differ in all-cause mortality.

No other well-conducted RCT has compared any other treatment with RP for clinically localized prostate cancer. One ongoing RCT—the Prostatectomy Intervention Versus Observation Trial (PIVOT) in the United States—will produce results in the future.\textsuperscript{90,91}

One observational study that used internal controls and data from the SEER program provided information on the effectiveness of RP relative to other treatments.\textsuperscript{92} For men with well-differentiated cancer, 10-year disease-specific survival did not differ between the RP group and age-matched radiation or watchful-waiting groups. Disease-specific survival was slightly higher for the RP group in men with moderately differentiated tumors (RP group, 87%; radiation group, 76%; watchful waiting group, 77%) and was much higher for men with poorly differentiated cancer (RP group, 67%; radiation group, 53%; watchful waiting group, 45%). Other cohort studies without controls have found similar survival rates after radical prostatectomy.\textsuperscript{33,93-97}

**Radiation Therapy**

Radiation therapy is the second most commonly used treatment for nonmetastatic prostate cancer and is the most common treatment for men ages 70 to 80 years.\textsuperscript{2} The two common types of radiation therapy reviewed here are external beam radiation therapy (EBRT) and brachytherapy, the insertion of radioactive pellets directly into prostate tissue.

No well-conducted RCT with clinical outcomes compares EBRT with any other therapy for clinically localized prostate cancer. In the large cohort study discussed earlier, 10-year disease-specific survival rates for men in the EBRT group were similar.
to those in the watchful-waiting group for men with well-differentiated and moderately differentiated tumors, but were higher for men with poorly differentiated cancer.92

Brachytherapy is most often used alone for men with well- or moderately differentiated intracapsular prostate cancer or in combination with EBRT for men with more aggressive cancer. No RCT with clinical outcomes compared brachytherapy with any other treatment for prostate cancer. Two observational studies with 100 or more patients with clinically localized prostate cancer found high survival for patients treated with radioactive gold or iodine seeds.98,99

Androgen Deprivation Therapy

The traditional approach to androgen deprivation therapy (ADT) has been surgical bilateral orchiectomy. A newer approach uses LHRH agonists (eg, goserelin or leuprolide), a group of drugs that stimulate the release of luteinizing hormone from the pituitary gland. Paradoxically, when used clinically, LHRH agonists result in a down regulation of pituitary receptors, thus markedly reducing the level of testosterone production to that of a castrated man. LHRH agonists have been used clinically since the late 1980s.

Two well-conducted RCTs compared clinical outcomes between men with clinically localized prostate cancer who were treated with ADT (with orchiectomy100 or estramustine101) and men treated with EBRT. ADT either increased overall survival100 or reduced clinical recurrence101; outcomes improved primarily among men who had lymph node involvement.

Four additional RCTs of ADT (with LHRH agonists) as an adjuvant to EBRT or RP for locally advanced prostate cancer found statistically significant improved overall survival (10% to 20% absolute difference) in men who received ADT.102-108 Another RCT of immediate versus deferred ADT (with orchiectomy or LHRH agonists) and no other treatment found improved survival (8% absolute difference) for the immediate therapy group in men who had new diagnosis of locally advanced prostate cancer.109

Watchful Waiting

The term watchful waiting implies that no treatment is given initially but that the patient is followed for evidence of progressive or symptomatic disease, for which treatment might be offered. Because the only well-conducted RCT that compares watchful waiting and more aggressive treatment examined men with prostate cancer detected clinically rather than by screening,48 the best information about the outcomes of watchful waiting comes from observational studies of men who, for various reasons, were not treated for prostate cancer. These studies also provide information about the natural history of the disease.

Four well-conducted retrospective cohort studies28,29,30,110,111 and one pooled analysis of six other cohort studies29 provide information about survival with untreated prostate cancer. Men with well-differentiated, clinically localized prostate cancer have excellent long-term survival, with little or no reduction in survival compared to similar men without prostate cancer. Men with poorly differentiated cancer have reduced survival. In one study, 10-year survival was 17% for men with poorly differentiated cancer and 47% in age-matched controls without prostate cancer.93

Because most prostate cancer detected by screening today is moderately differentiated, survival of men with this type of tumor is important to the debate about screening. On the standard histologic grading system for prostate cancer, these men have tumors with Gleason scores of 5 to 7. Gleason scores range from 2 to 10; lower scores indicate well-differentiated patterns and higher scores indicate more poorly differentiated tumors.

The most detailed analysis of men with untreated, clinically localized, moderately differentiated cancer found that 15-year prostate-cancer-specific survival rates ranged from 30% (Gleason score of 7 in men 50 to 59 years of age) to 94% (Gleason score of 5 in men 50 to 59 years of age).28,110 Men in their 70s had survival rates similar to those for men in their 50s for tumors with a Gleason score of 5 but much better survival for tumors with a Gleason score of 7 (58% vs. 30%). Because men in this study received their diagnoses in the pre-PSA era,
survival would probably be even better in similar men receiving diagnosis today given the “lead time” added by earlier detection.

Harms of Treatment

Because harms of treatment are experienced by the men themselves, and because men may have problems that are similar to treatment harms but are not attributable to treatment, we prioritized evidence that measured patients’ perceptions of their own function. For comparison, we used a non-treated group or the same men examined before and at least 12 months after treatment.

Radical Prostatectomy

Thirty-day mortality after RP ranges from 0.3% to 1% for most men and may be higher for men older than 80.112-116 The primary long-term adverse effects of RP include erectile dysfunction and urinary incontinence (Table 2). At least 20%, perhaps as many as 70%, of men have worsened sexual function as a result of RP.114,117-135 Fifteen percent to 50% of men who had an RP had some urinary problems 1 year later.112,114,117,121,123,126,127,131,132,135-138 Current evidence is mixed about the extent to which, outside of excellent academic centers,95,124 the newer nerve-sparing procedure reduces complication rates.

Radiation Therapy

Twenty percent to 45% of men with no erectile dysfunction and 2% to 16% of men with no urinary incontinence before EBRT developed dysfunction 12 to 24 months afterward (Table 2).117,123,126-131,135,137,139-146 Six percent to 25% of men who had no bowel dysfunction before EBRT reported marked problems 12 or more months afterward.117,123,125,126,130,131,135,137,140,142,144-147 The evidence is mixed about whether newer techniques, including 3-dimensional conformal EBRT, reduce the frequency of urinary or bowel side effects.

Compared with EBRT or RP, fewer high quality studies of the harms of brachytherapy have been completed. Our estimates are therefore less precise for this treatment. Among men who were potent before treatment, about 21% are impotent and 36% have decreased erectile function 3 years after brachytherapy.148,149 A majority of men will have distressing urinary symptoms in the first months after brachytherapy; from 6% to 12% will have such symptoms 1 year later. Up to 25% of men will have some lack of urinary control 12 months after brachytherapy.120,149-151 Approximately 18% of men will have diarrhea 1 year later;120 and 19% will have some persistent rectal bleeding 12 to 28 months later.152

Table 2. Harms of treatment*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Men with reduced sexual function</th>
<th>Men with urinary problems</th>
<th>Men with bowel problems</th>
<th>Men with other problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radical prostatectomy</td>
<td>20%–70%</td>
<td>15%–50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>External beam radiation therapy</td>
<td>20%–45%</td>
<td>2%–16%</td>
<td>6%–25%</td>
<td></td>
</tr>
<tr>
<td>Brachytherapy</td>
<td>36%†</td>
<td>6%–12%†</td>
<td>18%†</td>
<td>Breast swelling: 5%–25%</td>
</tr>
<tr>
<td>Androgen deprivation therapy</td>
<td>40%–70%</td>
<td></td>
<td></td>
<td>Hot flashes: 50%–60%</td>
</tr>
</tbody>
</table>

* Percentage of men treated who had side effects at least 12 months after treatment.
† These findings are less certain than other entries because they are based on less, or less good, evidence.
Androgen Deprivation Therapy

We focused on the harms of LHRH because the effectiveness studies we reviewed primarily used this type of androgen deprivation therapy. No study has examined reports from the same patients beginning before ADT and extending for at least one year. Our best information comes from 2 large national studies\textsuperscript{153-155} and a systematic review.\textsuperscript{156,157} Compared to untreated men, 40\% to 70\% of men who were sexually active before treatment were not sexually active afterward (Table 2). Five percent to 25\% of men had breast swelling, and 50\%-60\% had hot flashes. Mean scores on quality of life indices are lower for men treated with ADT.\textsuperscript{154} One RCT of LHRH adjuvant therapy found similar results.\textsuperscript{107} Potential long-term complications of LHRH therapy include lack of vitality, anemia and osteoporosis.\textsuperscript{155,158,159} The frequency and severity of these complications are not yet clear.

Quality of Life

Litwin et al. compared overall quality-of-life scores among control men and men with prostate cancer within treatment groups.\textsuperscript{129} Although they found the same differences in specific symptoms as noted earlier, they found no differences among groups (either among treatment groups or between men with and without prostate cancer) in overall quality of life.

Cost-effectiveness of Screening

Given the uncertainties about the existence and magnitude of benefits, the cost-effectiveness of screening for prostate cancer has been difficult to calculate. A 1993 decision analysis, which made optimistic assumptions about benefit from screening and early treatment, found little or no benefit for men with well-differentiated tumors.\textsuperscript{160} For men with moderately or poorly differentiated cancer, screening and early treatment could offer as much as 3.5 years’ improvement in quality-adjusted life expectancy, again using the most optimistic assumptions. Even with optimistic assumptions, men ages 75 years and older were not likely to benefit from screening and aggressive treatment. One major reason is that any benefits of screening are expected to accrue some years into the future, after many men of this age have died of some other condition. Two subsequent decision analyses have reached the same conclusions.\textsuperscript{161,162}

In 1995, Barry et al. published a cost-effectiveness analysis using favorable screening assumptions.\textsuperscript{163} The marginal cost-effectiveness of screening men age 65 years with PSA and DRE, without adjustment for life quality and without discounting benefits, was between $12,500 and $15,000 per life-year saved. Changing only a few assumptions, however, quickly increased the marginal cost-effectiveness ratio to above $100,000 per life-year saved. This ratio would be even less favorable if a decrement in quality of life associated with the harms of treatment were considered. In 1997, these investigators updated their model with newer data and further assumptions favorable to screening; findings were similar.\textsuperscript{164}

Discussion

PSA and, to a lesser extent, DRE can detect prostate cancer at an earlier stage than it could be detected clinically. A major problem in considering the utility of screening, however, is the heterogeneity of prostate cancer itself. The large discrepancy between prostate cancer diagnoses and deaths indicates that some and probably most cancers detected by screening are clinically unimportant. Because precise evidence regarding the prognosis of prostate cancer of various types is lacking, researchers have not been able to define the most appropriate target of screening, ie, the types of cancer that will cause clinical symptoms and death and that can be treated better if detected earlier.

The efficacy of various types of treatment for clinically localized prostate cancer, and especially for the type of localized prostate cancer detected by screening, is largely unknown. Although one RCT found that RP reduced prostate cancer mortality compared with watchful waiting among men with symptomatic localized cancer, the magnitude of any additional benefit of detection and earlier treatment due to screening is still unknown. We lack direct evidence that EBRT, brachytherapy, or ADT is effective for clinically localized cancer. Each treatment for prostate cancer is associated with various
potential harms, including sexual, urinary, and bowel dysfunction.

The costs of a screening program for prostate cancer are potentially high. If treatment is extremely efficacious, then the cost-effectiveness of screening men ages 50 to 69 years may be reasonable; if treatment is less efficacious, the results may be net harm and high costs. Assuming that any potential benefit to screening accrues only after some years, men with a life expectancy of fewer than 10 years are unlikely to benefit. Because prostate cancer incidence and mortality rates are higher among black men, beneficial screening could have a larger absolute benefit in this ethnic group than in white men. The same uncertainties about screening, however, would apply.

Two RCTs of screening are in progress. Because of the problem of screening in control groups, however, some experts fear that even these trials may not provide a definitive answer about screening efficacy. If these trials find a reduction in prostate cancer mortality, further work will be required to determine whether the benefits outweigh the harms and costs for individuals or as a general policy. Research can help by developing new screening and treatment approaches that minimize harms and costs. If the trials show no benefit, other approaches to disease control, such as chemoprevention, will be necessary. In the interim, the efficacy of screening for prostate cancer remains uncertain.

Grant support: This study was developed by the RTI-UNC Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (Contract No. 290-97-0011), Rockville, MD.

Acknowledgments: We acknowledge at AHRQ the continuing support of David Atkins, MD, MPH, Director of the Clinical Prevention Program for the U.S. Preventive Services Task Force. We especially want to thank our USPSTF liaisons: Cynthia D. Mulrow, MD, MPH, University of Texas at San Antonio, and Paul S. Frame, MD, Cohocton, New York, and other members of the USPSTF who provided guidance and constructive feedback.

The investigators deeply appreciate the considerable support and contributions of staff from the RTI-UNC Evidence-based Practice Center, including Audrina J. Bunton, BA, Linda Lux, MPA, and Sonya Sutton, BSPH, for substantive and editorial work on this systematic review, and Loraine Monroe, for superior secretarial assistance. We also appreciate Rainer Beck, MD, Paul Godley, MD, MPH, and Kenneth Fink, MD, MPH for assistance in reviewing articles. In addition, we are indebted to staff from the University of North Carolina at Chapel Hill and the Cecil G. Sheps Center for Health Services Research, including Timothy S. Carey, MD, MPH, Director of the Sheps Center and Co-Director of the RTI-UNC Evidence-based Practice Center; Anne Jackman, MSW; Lynn Whitener, MLS, DrPH.; and Carol Krasnov.

We also owe our thanks to the peer reviewers of the complete systematic evidence review for their constructive feedback and insightful suggestions: Michael Barry, MD, Medical Practices Evaluation Center, Boston, MA; John W. Feightner, MD, MSC, FCFP, Chair of the Canadian Task Force on Preventive Health Care, London, Ontario; Theodore Ganiats, MD, University of California at San Diego, La Jolla, CA; Marc B. Garnick, MD, Beth Israel Deaconess Medical Center, Boston, MA; Richard Hoffman, MD, MPH, Albuquerque, VA Medical Center, Albuquerque, NM; Barnett Kramer, MD, National Institutes of Health, Bethesda, MD; David Lush, MD, Medical College of Pennsylvania, Philadelphia, PA; Curtis Mettlin, MD, Roswell Park Cancer Institute, Buffalo, NY; and Ian M. Thompson, Jr, MD, University of Texas Health Sciences Center at San Antonio, San Antonio, TX.
References


52. Catalona WJ, Smith DS, Ornstein DK. Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/mL and benign prostate examination. Enhancement of specificity with free PSA measurements. JAMA. 1997;277:1452-1455.


63. Catalona WJ, Richie JP, Ahmann FR, et al. Comparison of digital rectal examination and serum prostate specific antigen in the early detection of


105. Pilepich MV, BuzydloWSKI JW, John MJ, Rubin P, McGowan DG, Marcial VA. Phase II trial of hormonal cytoreduction with megestrol and diethylstilbestrol in conjunction with radiotherapy for carcinoma of the prostate: outcome results of


125. Helgason AR, Adolfsen J, Dickman P, Arver S, Fredrikson M, Steineck G. Factors associated with


146. Widmark A, Fransson P, Tavelin B. Self-assessment questionnaire for evaluating urinary and intestinal


Appendix
Screening for Prostate Cancer: An Update of the Evidence

Methods
This appendix documents procedures that the RTI-UNC Evidence-based Practice Center (EPC) staff used to develop this report on screening for prostate cancer. During preparation of the evidence report, we collaborated with two current members of the U.S. Preventive Services Task Force (USPSTF) who served as liaisons to the EPC topic team. We first document the analytic framework and key questions developed at the beginning of the review. We then describe the inclusion and exclusion criteria for admissible evidence, our strategy for literature search and synthesis, and our approach to developing the final summary of the evidence.

Analytic Framework and Key Questions
The analytic framework (Figure 1) describes the relationship between screening and treating patients in a clinical setting and reduced morbidity or mortality from prostate cancer. The arrows with superscripts in the analytic framework represent steps in the chain of logic connecting screening with reduced morbidity and/or mortality from prostate cancer; the superscripts refer specifically to 9 key questions that guided our literature searches and synthesis of the evidence. We examined 1 overarching question (Key Question 1, linking screening and ultimate health outcomes) and 8 additional questions pertaining to specific links in the analytic framework. The key questions were as follows:

Key Question 1: What are the health outcomes (both type and magnitude) of screening a defined population for prostate cancer compared to not screening?

Key Question 2: What is the yield of screening for prostate cancer (ie, accuracy and reliability of screening tests, prevalence of undetected cancer in various populations)?

Key Questions 3-6: What are the health outcomes associated with treating clinically localized prostate cancer with radical prostatectomy, external beam radiation therapy or brachytherapy, androgen deprivation, or watchful waiting?

Key Question 7: What harms are associated with the treatments of clinically localized prostate cancer?

Key Question 8: What costs are associated with screening for and early treatment of prostate cancer? Have studies modeled the potential benefits of screening? What is the cost-effectiveness of screening for prostate cancer?

Key Question 9: What harms are associated with screening for prostate cancer?

Because we found little evidence about Key Question 9, the harms of screening (one article with inconclusive results), we did not discuss this issue in the article submitted to the Annals of Internal Medicine.

Eligibility Criteria for Admissible Evidence
The EPC staff and Task Force liaisons developed eligibility criteria for selecting the evidence relevant to answer the key questions (see Table 1). We first searched for evidence from randomized controlled trials (RCTs) for the efficacy of screening. Because we found no well-conducted and well-analyzed RCT of screening, we then examined case-control and ecologic evidence regarding the overarching key question (Key Question 1).
For Key Question 2, concerning the operating characteristics of screening tests, we examined well-conducted systematic reviews and individual studies that started with a primary care or unselected population without prostate cancer and that compared the findings of one or more screening tests with an adequate reference standard. We also examined evidence of the yield of screening from well-conducted screening programs. For Key Questions 3 through 6, concerning the effectiveness of various therapies, we required evidence from RCTs. For Key Questions 7 and 9, concerning the harms of screening or treatment, we required either RCTs or well-controlled studies that included patient reports and the use of a valid measurement instrument. Finally, for Key Question 8, we searched for evidence of the costs and cost-effectiveness of screening, including models of potential benefits, that considered all appropriate costs and estimates of effectiveness supported by reasonable assumptions based on good evidence.

**Literature Search Strategy and Synthesis**

The analytic framework and key questions guided our literature searches. We examined the critical literature described in the review by the USPSTF (published in 1996) and searched the reference lists of systematic reviews (including Cochrane Library reviews) published since 1993. We then used our eligibility criteria to develop search terms and searched the MEDLINE database for relevant articles concerning humans in the English language published between January 1, 1994, and September 15, 2002. We especially looked for articles involving patients whose experience is clearly generalizable to a primary care U.S. population.

The search strategy and results are given in Table 1 and Figure 2. All searches started with the term “prostate neoplasm” and then proceeded by adding further terms as shown in Table 1.
### Appendix Table 1. Eligibility criteria, search strategy, and results of searches

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Inclusion Criteria</th>
<th>Search Terms</th>
<th>Number of Articles</th>
<th>Identified for Abstract Review</th>
<th>Retained for Full Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Efficacy of screening in reducing mortality from prostate cancer</td>
<td>- RCT or case-control&lt;br&gt;- Surveillance (ecologic) study of PC incidence, morbidity, or mortality over time</td>
<td>Prostate neoplasms&lt;br&gt;Mass screening&lt;br&gt;RCT&lt;br&gt;Case-control&lt;br&gt;Incidence&lt;br&gt;Mortality&lt;br&gt;Trends&lt;br&gt;Surveillance</td>
<td>1.399</td>
<td>100 RCT:1&lt;br&gt;Case-control: 2</td>
<td>Ecologic: 15</td>
</tr>
<tr>
<td>2. Yield of screening tests</td>
<td>- Unselected population without PC&lt;br&gt;- Screening test used for all&lt;br&gt;- Result of screening test compared with a valid reference standard</td>
<td>Prostate neoplasms&lt;br&gt;Mass screening&lt;br&gt;DRE, PSA&lt;br&gt;Diagnosis&lt;br&gt;Sensitivity/Specificity&lt;br&gt;Predictive value&lt;br&gt;Reproducibility&lt;br&gt;Screening programs</td>
<td>1,905</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>3-6. Health outcomes of treatment</td>
<td>- RCT&lt;br&gt;- Follow-up at least 2 years&lt;br&gt;- At least 75% of patients followed&lt;br&gt;- Health outcomes</td>
<td>Prostate neoplasms&lt;br&gt;Therapeutics&lt;br&gt;Treatment&lt;br&gt;Surgery&lt;br&gt;Prostatectomy&lt;br&gt;Radiation&lt;br&gt;Brachytherapy&lt;br&gt;RCT</td>
<td>656</td>
<td>KQ3: 1&lt;br&gt;KQ4: 0&lt;br&gt;KQ5: 2&lt;br&gt;KQ6: 1</td>
<td></td>
</tr>
<tr>
<td>7. Harms of treatment</td>
<td>- Unselected population with PC&lt;br&gt;- Treated group compared with valid comparison group&lt;br&gt;- Either randomized or adjustment for confounders&lt;br&gt;- Not metastatic cancer&lt;br&gt;- Valid measures of harms&lt;br&gt;- At least 75% of patients followed&lt;br&gt;- At least one year follow-up</td>
<td>Prostate neoplasms&lt;br&gt;Therapeutics&lt;br&gt;Treatment&lt;br&gt;Surgery&lt;br&gt;Prostatectomy&lt;br&gt;Radiation&lt;br&gt;Adverse effects&lt;br&gt;Side effects&lt;br&gt;Impotence&lt;br&gt;Urinary incontinence&lt;br&gt;Quality of life</td>
<td>923</td>
<td>923</td>
<td>32</td>
</tr>
<tr>
<td>9. Harms of screening</td>
<td>- Unselected population&lt;br&gt;- Screened group compared with unscreened group&lt;br&gt;- Randomized trial or adjustment for confounders&lt;br&gt;- Reliable measure of adverse effects</td>
<td>Prostate neoplasms&lt;br&gt;Mass screening&lt;br&gt;Adverse effects&lt;br&gt;Anxiety, depression&lt;br&gt;Labeling&lt;br&gt;Quality of life</td>
<td>94</td>
<td>94</td>
<td>1</td>
</tr>
</tbody>
</table>

**Note:** DRE indicates digital rectal examination; KQ, key question; PSA, prostate-specific antigen; RCT, randomized controlled trial.
The first author reviewed abstracts of all articles found in the searches to determine which met eligibility criteria. Other EPC authors of the full systematic evidence review (SER) reviewed all abstracts excluded by the first reviewer. We retrieved the full text of all articles not excluded by both reviewers (see next to last column in Table 1).

One reviewer then examined the full text of all retrieved articles against the eligibility criteria and discussed all excluded articles with one of the other reviewers. We included any article that either reviewer judged had met eligibility criteria (see last column in Table 1). Three of the authors of the SER then divided the articles and abstracted data from them, entering the relevant data into pre-designed evidence tables (see Appendix B to the SER, Screening for Prostate Cancer, on AHRQ Web site [www.ahrq.gov]). The abstracting author also graded the articles using the criteria established by the Methods Work Group of the USPSTF.5 The first author read all articles, checked the grading, and discussed the crucial ones with a second reviewer. The authors also discussed key articles with the Task Force liaisons.

---

**Appendix Figure 2. Selection of articles for screening for prostate cancer**

**Key Question 1: Efficacy of Screening (Direct Evidence)**

**RCTs and Case-Control Studies**

Articles from MEDLINE and other searches (n=100) → Articles excluded: not RCTs or case-control studies (n=97) → Articles retrieved for detailed evaluation (n=3)

**Key Question 1: Efficacy of Screening (Ecologic Studies)**

Articles from MEDLINE and other searches (n=1,399) → Articles excluded: no high quality incidence data; no correlation with screening rates; no high quality mortality data (n=1,384) → Articles retrieved for detailed evaluation (n=15)

(Continued on p. 33)
Development of the Final Systematic Evidence Review

We presented an initial work plan, including a provisional analytic framework and key questions, to the entire Task Force in September 2000; we also presented interim reports on results of the literature search and the early results of the synthesis of information in December 2000 and March 2001. A draft of the SER was submitted for broad-based external peer review in May 2001; the peer review involved individual experts in the field, representatives of relevant professional organizations, and representatives of organizations and federal agencies that serve as liaisons to the USPSTF. We revised the evidence review as appropriate after receiving peer review comments. The Task Force reviewed all information and voted on a recommendation in June 2001, revising the Rationale and Recommendation Statement in the spring of 2002 after review by involved professional associations and agencies. We then updated the searches, finalized the review, and shortened it for publication.

Appendix Figure 2. Selection of articles for screening for prostate cancer (cont.)

Key Question 2: Yield of Screening Tests

Articles from MEDLINE and other searches (n=1,905)

Articles excluded: not population-based, no valid gold standard, not prostate cancer (n=1,870)

Articles retrieved for detailed evaluation (n=35)

Key Questions 3-6: Health Outcomes of Treatment

Articles from MEDLINE and other searches (n=656)

Articles excluded: not RTC, not prostate cancer, inappropriate population (n=653)

Articles retrieved for detailed evaluation (n=3)

(Continued on p. 34)
Key Question 7: Harms of Treatment

Articles from MEDLINE and other searches (n=923)

Articles excluded: no valid self-reports, no comparison with untreated men or with treated men before treatment, selected population (n=891)

Articles retrieved for detailed evaluation (n=32)

Key Question 8: Costs/Cost-Effectiveness of Screening

Articles from MEDLINE and other searches (n=94)

Articles excluded: not prostate cancer screening, selected population, no description of methods (n=82)

Articles retrieved for detailed evaluation (n=2)

Key Question 9: Harms of Screening

Articles from MEDLINE and other searches (n=93)

Articles excluded: not prostate cancer screening, no valid measurement of psychological or physical harm (n=93)

Articles retrieved for detailed evaluation (n=1)