Evidence Report/Technology Assessment
Number 204

An Evidence Review of Active Surveillance in Men With Localized Prostate Cancer

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Evidence-Based Practice

Evidence Report/Technology Assessment

Number 204

An Evidence Review of Active Surveillance in Men With Localized Prostate Cancer

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10055-I, Task Order No. 10

Prepared by:
Tufts Evidence-based Practice Center, Tufts Medical Center
Boston, MA

Investigators:
Stanley Ip, M.D., Project Lead
Issa J. Dahabreh, M.D., M.S.
Mei Chung, Ph.D., M.P.H.
Winifred W. Yu, Ph.D., M.S.
Ethan M. Balk, M.D., M.P.H.
Ramon C. Iovin, Ph.D., Editor
Paul Mathew, M.D., Clinical Content Expert
Tony Luongo, M.D., Clinical Content Expert
Tomas Dvorak, M.D., Clinical Content Expert
Joseph Lau, M.D.

AHRQ Publication No. 12-E003-EF
December 2011

This report is based on research conducted by the Tufts Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA 290-2007-10055-I). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care researchers and funders of research make well-informed decisions in designing and funding research and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of scientific judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical research and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances.

This report may be used, in whole or in part, as the basis for research design or funding opportunity announcements. AHRQ or the U.S. Department of Health and Human Services endorsement of such derivative products or actions may not be stated or implied.

Persons using assistive technology may not be able to fully access information in this report. For assistance, contact EffectiveHealthCare@ahrq.hhs.gov.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Ip S, Dahabreh IJ, Chung M, Yu WW, Balk EM, Iovin RC, Mathew P, Luongo T, Dvorak T, Lau J. An Evidence Review of Active Surveillance in Men With Localized Prostate Cancer. Evidence Report/Technology Assessment No. 204. (Prepared by Tufts Evidence-based Practice Center under Contract No. HHSA 290-2007-10055-I.) AHRQ Publication No. 12-E003-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2011. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. This report was requested and funded by the National Institutes of Health (NIH) Office of Medical Applications of Research (OMAR). The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ—based on deliberations by the Planning Committee convened by OMAR, the National Cancer Institute, the Centers for Disease Control and Prevention, the National Institute on Aging, and the Department of Veterans Affairs—and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review prior to their release as a final report.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

We welcome written comments on this evidence report. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Jennifer Miller Croswell, M.D., M.P.H.
Acting Director
Office of Medical Applications of Research
Office of the Director
National Institutes of Health

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

William Lawrence, M.D., M.S.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

We would like to acknowledge the help and constructive comments provided by members of the Technical Expert Panel and peer reviewers in bringing this report to fruition.

Technical Expert Panel

Hans-Olov Adami, M.D., Ph.D.
Harvard School of Public Health
Boston, MA

Patricia Ganz, M.D.
University of California Los Angeles School of Public Health
Los Angeles, CA

Paul Godley, M.D., Ph.D.
University of North Carolina at Chapel Hill
Chapel Hill, NC

Robert B. Hicks, Esq.
Patient Representative

Deborah Kuban, M.D.
The University of Texas MD Anderson Cancer Center
Houston, TX

Daniel W. Lin, M.D.
University of Washington
Seattle, WA

Daniel A. Ollendorf, M.P.H.
Institute for Clinical and Economic Review
Boston, MA

Peter T. Scardino, M.D.
Memorial Sloan-Kettering Cancer Center
New York, NY

Invited Peer Reviewers

Michael Barry, Ph.D.
Massachusetts General Hospital
Boston, MA

Justin Bekelman, M.D.
Hospital of the University of Pennsylvania
Philadelphia, PA

Matthew Cooperberg, M.D., M.P.H.
Hospital of the University of California, San Francisco
San Francisco, CA

Julia Hayes, M.D.
Dana-Farber Cancer Institute
Boston, MA

Eric Klein, M.D.
Glickman Urological & Kidney Institute, Cleveland Clinic
Cleveland, OH

Steven Pearson, M.D.
Institute for Clinical and Economic Review
Boston, MA

Mark Soloway, M.D.
University of Miami School of Medicine
Miami, FL

An Evidence Review of Active Surveillance in Men With Localized Prostate Cancer

Structured Abstract

Background. Radical prostatectomy and radiation therapy for prostate cancer have side effects and unclear survival benefits for early stage and low-risk disease. Prostate cancer often has an indolent natural history, making observational management strategies potentially appealing.

Purpose. To systematically review the role of active surveillance for triggers to begin curative treatment in men with low-risk prostate cancer. Key Questions address changes in prostate cancer characteristics over time, definitions of active surveillance and other observational strategies, factors affecting the offer of, acceptance of, and adherence to active surveillance, the comparative effectiveness of active surveillance with curative treatments, and research gaps.

Data sources. MEDLINE®, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and existing systematic reviews, evidence reports, and economic evaluations.

Study selection. Randomized controlled trials and nonrandomized comparative studies of treatments, multivariable association studies, and studies of temporal trends in prostate cancer natural history. Only published, peer-reviewed, English-language articles were selected based on predetermined eligibility criteria.

Data extraction. A standardized protocol was used to extract details on design, diagnoses, interventions, predictive factors, outcomes, and study validity.

Data synthesis. In total, 80 studies provided information on epidemiologic trends; 56 on definitions of active surveillance; 42 on factors affecting the offer of, acceptance of, or adherence to observational management strategies; and 26 on comparative effectiveness. Increased diagnosis of early-stage prostate cancer due to prostate-specific antigen (PSA) testing, led to an increase in prostate cancer incidence from the mid-1980s to the mid-1990s. The prostate cancer-specific mortality rate decreased for all age groups from the early-1990s to 1999. Currently, patients are diagnosed with earlier stage and lower risk prostate cancers compared to the pre-PSA era. Over time, a lower proportion of men received observational management versus active treatment, even among those with low-risk disease. There was no standardized definition of active surveillance. Sixteen cohorts used different monitoring protocols, all with different combinations of periodic digital rectal examination, PSA testing, rebiopsy, and/or imaging findings. Predictors that a patient received no initial active treatment generally included older age, presence of comorbidities, lower Gleason score, lower tumor stage, lower diagnostic PSA, and lower disease progression risk group. No trial provided results comparing men with localized disease on active surveillance with surgery or radiation therapy.

Limitations. Because of the nonstandardized usages of the terms “active surveillance” and “watchful waiting” and their intended and often mixed (both curative and palliative) treatment objectives, it was difficult to determine which study patients received active monitoring for

triggers indicative of curative treatment and which observation for clinical symptoms indicative of palliative treatment.

Conclusions. More men are being diagnosed with early stage prostate cancer. Whether active monitoring with a curative intent is an appropriate option for these men remains unclear. A standard, universally agreed-upon definition of active surveillance that clearly distinguishes it from watchful waiting and other observational management strategies is needed to help clarify scientific discourse on this topic. Ongoing clinical trials may provide information on the comparative effectiveness of active surveillance compared to immediate active treatment, but will require long term followup.

Contents

Executive Summary ES-1
Introduction 1
Key Questions 2
Methods 5
AHRQ Task Order Officer 5
External Expert Input 5
Key Questions 5
Analytic Framework 5
Literature Searches 6
Study Selection and Eligibility Criteria 7
Systematic Reviews, Evidence Reports, Economic Evaluations 7
Primary Research Studies 8
Data Extraction and Summaries 11
Primary Research Studies 11
Systematic Reviews 12
Quality Assessment 12
Primary Research Studies 12
Systematic Reviews 13
Data Synthesis and Presentation 13
Summary Tables 14
Grading the Body of Evidence for Key Question 4 14
Peer Review and Public Commentary 15
Results 16
Key Question 1. How have the patient population and the natural history of prostate cancer diagnosed in the United States changed in the last 30 years? 17
Trends in Prostate Cancer Incidence 20
Trends in Prostate Cancer Mortality and Survival Rates 22
Patient, Tumor, and System-Level Characteristics at Diagnosis 24
Trends in Treatment Patterns 28
Summary 28
Key Question 2. How are active surveillance and other observational management strategies defined? 29
Protocols With Curative Intent 30
Observational Management Strategies With Palliative Intent 47
Observational Management Strategies With Unclear Treatment Intent 58
Summary 58
Key Question 3. What factors affect the offer of, acceptance of, and adherence to active surveillance? 63
Physician Factors (Appendix Tables C3.1–3.3) 65
Clinical Factors 67
Patient Factors 68
Other Factors That Could Affect the Offer of, Acceptance of, or Adherence to AS 70
Summary 74

Key Question 4. What are the comparative short- and long-term outcomes of active surveillance versus immediate treatment with curative intent for localized prostate cancer? 75
Clinical Outcomes 75
Costs 81
Summary 83
Key Question 5. What are the research needs regarding active surveillance (or watchful waiting) in localized prostate cancer? 84
Key Question 1. Patient Population and Natural History Changes in Last 30 Years 85
Key Question 2. Definition of Active Surveillance 86
Key Question 3. Factors That Affect Offer of, Acceptance of, and Adherence to AS 87
Key Question 4. Active Surveillance Versus Immediate Curative Treatment 88
Discussion 90
References 93
Abbreviations and Acronyms 106

Tables

Table 1. Unique AS Cohorts 31
Table 2. Cohorts That Used Gleason Score and PSA as Part of Patient Eligibility Criteria for AS 33
Table 3. Eligibility Criteria for Enrollment in Protocols With Curative Intent in Chronological Order of Starting Enrollment Year 34
Table 4. Cohorts That Used PSA (ng/mL) or PSA Kinetics as Part of Followup Protocol for AS 40
Table 5. Monitoring Criteria in Protocols With Curative Intent in Chronological Order of Starting Enrollment Year 41
Table 6. Unique 13 Cohorts of Observational Management Strategies With Palliative Intent 47
Table 7. Eligibility Criteria for Enrollment in Protocols With Palliative Intent in Chronological Order of Starting Enrollment Year 50
Table 8. Monitoring Criteria in Protocols of Observational Management Strategies With Palliative Intent in Chronological Order of Starting Enrollment Year 53
Table 9. Cohorts That Did Not Report Triggers for Treatment of Prostate Cancer 58
Table 10. Protocols That Did Not Report Information on Triggers for Intervention in Chronological Order of Starting Enrollment Year 60
Table 11. Monitoring Parameters in Cohorts That Did Not Report Information on Triggers for Intervention in Chronological Order of Starting Enrollment Year 61

Figures

Figure A. Analytic Framework That Depicts the Five Key Questions (KQs) That Examine the Role of Active Surveillance in the Management of Men With Clinically Localized Prostate Cancer ES-5
Figure 1. Analytic Framework That Depicts the Five Key Questions (KQs) That Examine the Role of Active Surveillance in the Management of Men With Clinically Localized Prostate Cancer 6
Figure 2. Literature Flow 16
Figure 3. Years Covered and Databases Utilized by Studies Considered Eligible for Key Question 1 19

Figure 4. Age-Adjusted SEER Incidence Rates for Prostate Cancer (1975–2008) 20
Figure 5. Followup Frequencies of 16 Unique AS Cohorts 39
Figure 6. Followup Frequencies of 13 Unique Cohorts of Observational Management Strategies With Palliative Intent 56
Figure 7. Evidence Map for Key Question 3 64

Appendixes

Appendix A. Search Strategies
Appendix B. Ongoing Randomized Studies Comparing Observational Management Strategies With Active Treatment Strategies for the Treatment of Clinically Localized Disease
Appendix C. Appendix Tables for Key Questions 1–4

Executive Summary

Background[a]

In 2011, more than 240,000 men are projected to be diagnosed with prostate cancer, and 33,000 are projected to die from the disease in the United States. In the United States, most cases of prostate cancer are detected via prostate-specific antigen (PSA) screening. The cancer is usually localized, and most tumors have low histological grades and low Gleason scores. Indeed, more than half of prostate cancers detected by PSA screening are expected to be early-stage, low-risk tumors. Such cancers are an infrequent cause of death, and those affected are more likely to die of unrelated causes.

A number of immediate active treatment options are available for localized prostate cancer. Most commonly, radical prostatectomy (RP) or radiation therapy (RT), with or without androgen deprivation therapy (ADT), are offered with curative intent. However, the clinical benefit of immediate therapy with curative intent has not yet been demonstrated for localized prostate cancer in a PSA-screened population. It is likely that a large number of men are receiving active treatment with curative intent without much likelihood of obtaining any clinical benefit due to the slow progression of many prostate tumors. Both surgical and radiation treatments result in significant short- and long-term adverse events, including impotence, urinary dysfunction, and other complications. Thus, determination of the appropriate management strategy for early-stage, low-risk prostate cancer is an important public health concern.

Active surveillance (AS) and watchful waiting (WW) are two observational followup strategies that forgo immediate therapy in patients with prostate cancer, with the goal of minimizing the morbidities and costs of immediate active treatment for men who may never develop cancer-related symptoms or who are interested in palliative treatments only. AS is curative in intent, and WW is palliative. AS is appropriate in men with disease believed to be indolent and therefore may not require therapy. Because prediction tools are imperfect, these men are monitored closely and treated with curative intent at signs of progression or patient choice. In this way, the considerable adverse effects of treatment are at best avoided, and at least deferred. This approach is to be distinguished from men for whom treatment is deemed inappropriate because of comorbidity; for these men, WW is generally considered, as it offers the option of palliative therapy upon symptomatic disease progression. AS often entails a multifactorial followup of patients—monitoring of PSA values, digital rectal examinations (DRE), prostate imaging, and periodic prostate biopsies—while WW is commonly a relatively passive strategy—with interventions triggered by symptoms. However, there is a continuum of aggressiveness of followup for both AS and WW, as practiced in the community. It should be noted that even though the two terms are used commonly in the scientific literature, the attended intents (curative vs. palliative) of these approaches are not always made clear. Furthermore, many analyses or databases combine AS, WW, and noncurative interventions like primary ADT in their analyses, making it impossible to ferret out issues specifically related to AS.

Immediate active treatment has tradeoffs, including the harms of short- and long-term complications from curative treatments and the benefits of potential reductions in long-term morbidity and mortality. Thus, AS and other observational management strategies may be considered by men who are more interested in avoiding the risks of curative treatment.

[a] Please refer to the reference list in the full report for a full documentation of statements contained in the Executive Summary.

Therefore, it is important to clarify the appropriate eligibility criteria and followup protocols for these observational strategies that could minimize both unnecessary early curative treatments and avoidable prostate cancer symptoms and deaths. Of course, this assumes that AS is as effective as (or no worse than) immediate curative treatments in an appropriate subgroup of men diagnosed with prostate cancer. This, however, remains to be proven. Furthermore, some men may be uncomfortable with observational management and feel a strong need to "do something," and thus AS may be rarely offered, chosen, or adhered to. Therefore, the factors affecting these actions also warrant further investigation.

The National Cancer Institute and the Centers for Disease Control and Prevention are sponsoring a National Institutes of Health (NIH) State-of-the-Science Conference in December 2011 to examine these and other essential issues regarding the role of AS (as opposed to immediate curative intent therapy) in the management of early-stage, low-risk prostate cancer. The NIH has tasked the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) Program to provide the present review of evidence for use in this conference.

Objectives

The objective of this report is to summarize the existing literature regarding the role of AS in the management of early-stage, low-risk prostate cancer. Both the report and the corresponding NIH State-of-the-Science conference are a part of the NIH Consensus Development Program, the purpose of which is to evaluate the scientific evidence on a particular topic and develop a consensus statement that advances research in that area. This statement is developed by an independent panel that is assembled for the conference. The panel on AS will hear the scientific data, including the findings of the present evidence review, and will then use that information to compose their statement. Additional information about the NIH Consensus Development Program can be found at: http://consensus.nih.gov/.

The Conference planning committee crafted the Key Questions to be addressed at the conference, and the EPC was charged with systematically reviewing the relevant literature to address them. Key Question 1 pertains to temporal trends in the natural history of prostate cancer in the United States. Key Question 2 relates to the definitions of observational management strategies (i.e., those involving no active treatment) for prostate cancer used in the published literature. Key Question 3 relates to the factors that influence the offer or acceptance of or adherence to AS. Key Question 4 pertains to the comparative effectiveness of AS versus active treatments for localized prostate cancer. Key Question 5 addresses recommendations for future research on observational management strategies for localized prostate cancer. The exact wordings of the Key Questions provided to the EPC for systematic review are as follows:

Key Questions

1. How have the patient population and the natural history of prostate cancer diagnosed in the United States changed in the last 30 years?
    a. Patient Characteristics
        i. Age
        ii. Comorbidity
        iii. Race/ethnicity
    b. Tumor Characteristics
        i. Stage

ii. Tumor volume
iii. Gleason score
iv. PSA

c. Diagnostic Strategies
i. Biopsy Frequency
ii. # of cores
iii. Histopathologic grading changes

d. System Characteristics
i. Differences in geographical access

2. How are active surveillance and other observational management strategies defined?
a. Common Metrics
i. Age
ii. Gleason
iii. # cores
iv. % cores
v. PSA (velocity, doubling time)
vi. Imaging
vii. Behavioral indicators

b. Followup Protocols
i. Gleason
ii. # cores
iii. % cores
iv. PSA
v. Imaging
vi. Behavioral indicators

3. What factors affect the offer of, acceptance, and adherence to active surveillance?
a. Physician Factors
i. Primary care
ii. Diagnosing physician
iii. Consultant – second opinion
iv. Clinical factors

b. Patient Factors
i. Family involvement
ii. Personal preferences
iii. Risk perceptions
iv. Family history
v. Social support

c. Delivery System
i. Economic incentives and disincentives
1. Insurance Type (HMO, military, private)
2. Availability of technology

ii. Geographic location
1. Small area variation
2. Regional variation
3. Urban vs. rural

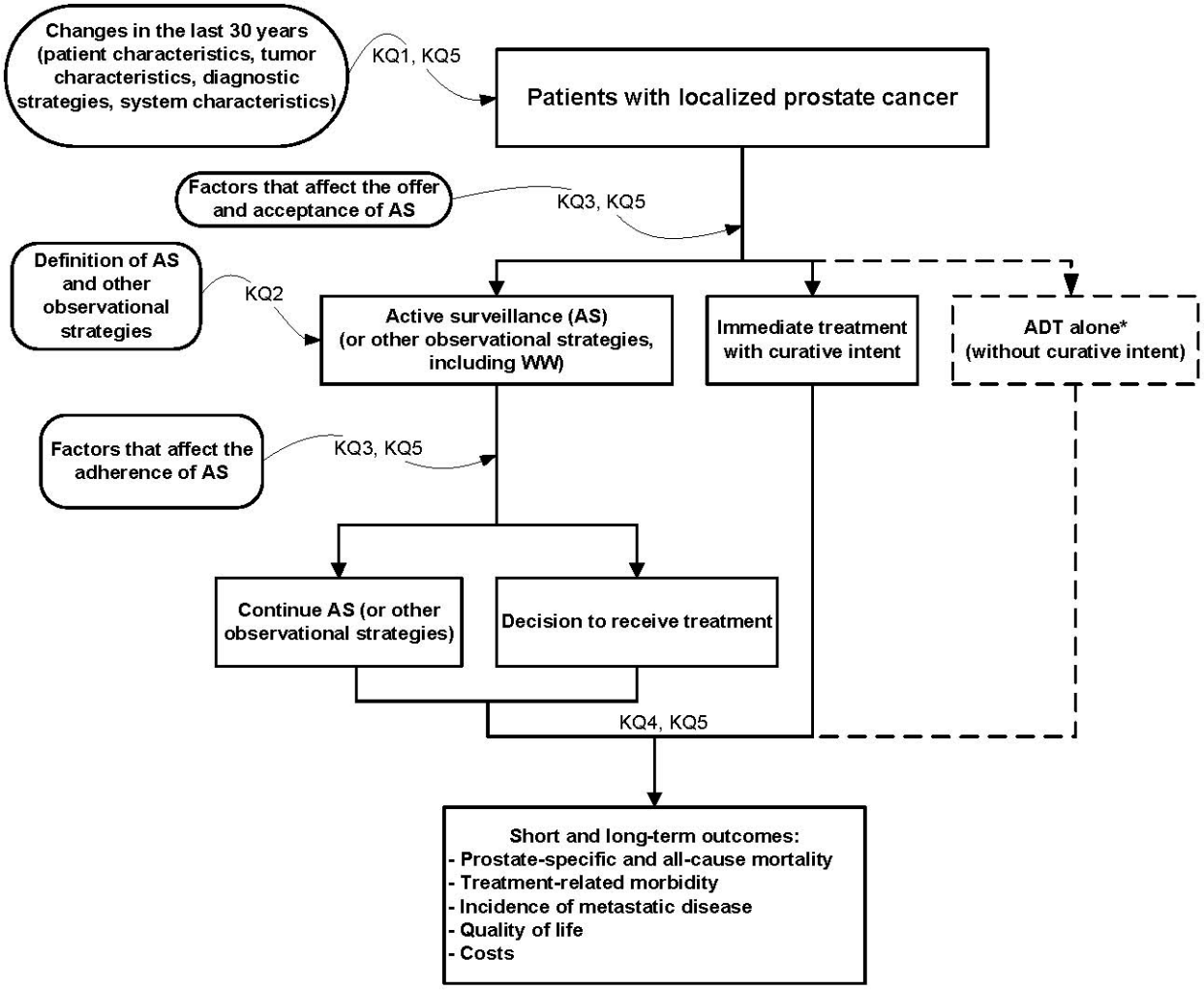
iii. Academic centers vs. private practice

    d. Communication Strategies
        i. Risk assessment, predictive models
        ii. Decisionmaking tools and aids
4. What are the comparative short- and long-term outcomes of active surveillance versus immediate treatment with curative intent for localized prostate cancer?
    a. Prostate-specific and all cause mortality
    b. Morbidity of primary treatment decision
    c. Incidence of metastatic disease
    d. Quality of life
    e. Costs
5. What are the research needs regarding active surveillance (or watchful waiting) in localized prostate cancer?

Analytic Framework

To guide this systematic review and facilitate the interpretation of Key Questions, we developed an analytic framework (Figure A) that depicts the logical progression and interconnection of all five Key Questions for this report.

Figure A. Analytic framework that depicts the five Key Questions that examine the role of active surveillance in the management of men with clinically localized prostate cancer



ADT = androgen deprivation therapy; AS = active surveillance; KQ = Key Question.
* ADT alone is a management strategy that is not reviewed in this report.

Methods

The EPC convened a group of experts in the epidemiology and treatment of prostate cancer to form a Technical Expert Panel, which provided clinical and methodological expertise in interpreting the Key Questions, identifying important issues, and defining parameters for the review of evidence. In addition, input from these experts was sought when questions arose regarding the scope of the review.

Literature Searches, Eligibility Criteria, and Screening

Multiple literature searches were performed in MEDLINE from inception to August 2011. We searched for recent systematic reviews, and subsequently conducted separate but overlapping searches for each of the first four Key Questions. We used search terms related to prostate cancer, active surveillance, watchful waiting, expectant management, and other related

management strategies. We also searched for studies of specific databases, including SEER (Surveillance Epidemiology and End Results) and CaPSURE (Cancer of the Prostate Strategic Urologic Research Endeavor). For Key Question 4, we relied on previous systematic reviews on prostate cancer conducted for the AHRQ EPC program. Searches were supplemented with studies recommended to us by the Technical Expert Panel, reference lists of eligible primary studies and relevant review articles, and targeted searches for economic evaluations. We did not include unpublished data.

Below are the study eligibility criteria we used for the first four Key Questions (no specific literature search was performed for Key Question 5):

Key Question 1. Studies of large U.S.-based databases of patients with prostate cancer with time-trend data (reporting changes over a range of years) between 1980 and 2011. Studies must have had a sample size of at least 1000 patients. We also reviewed the latest version of the American Cancer Society Cancer Statistics report[1], a recent SEER Survival Monograph[2], and data available on the SEER Web site.[b]

Key Question 2. Studies of any design that reported protocols and management strategies for patients receiving observational management (i.e., no immediate curative treatment). We included both studies where the goal of observation was to identify disease progression indicative of the need for curative treatments, and studies where the goal of observation was to determine the need for palliative treatments.

Key Question 3. Three types of studies were included. Firstly, we included studies that used quantitative methods to analyze databases or cohorts of patients to elucidate predictors of the offer or acceptance of or adherence to observational management strategies (including AS and WW). We excluded studies that analyzed ADT together with observational management strategies. We required multivariable analyses adjusting for a minimum of age and tumor stage (if the analysis was not limited to localized cancer) or using a propensity score. Secondly, we included studies using qualitative research methods (e.g., focus groups or surveys) to obtain information on factors that affect the offer or acceptance of or adherence to AS or WW. Eligible studies must have used a predefined approach to collect information. Thirdly, we also searched for experimental studies evaluating the effect of tools, such as decision aids, on the offer or acceptance of or adherence to AS (however, no such studies were found).

Key Question 4. We included randomized and nonrandomized, prospective or retrospective longitudinal comparative studies performed in a multicenter setting. Nonrandomized studies must have used multivariable or other methods to adjust for possible confounding, specifically for age and tumor stage, to warrant inclusion. The population of interest was men with clinically localized prostate cancer (T1-T2), without known lymph nodes (N0-X) or metastases (M0-X). No more than 20 percent of the study sample could exhibit more advanced disease. Studies had to compare observational management strategies (without ADT) to active treatment, including RP, external beam RT (EBRT), or brachytherapy (BT), all with or without ADT. However, ADT monotherapy was not considered an active treatment. Outcomes of interest included: prostate-cancer mortality, all-cause mortality, morbidity of primary treatment, metastatic disease, quality of life, and costs.

[b] Available at http://seer.cancer.gov/faststats/; last accessed September 30, 2011.

All five EPC team members participated in screening and selecting studies. An iterative screening process was used for training and to ensure consistency in application of eligibility criteria. Abstracts were screened once. A very low threshold was used to mark a study as of possible interest. During full-text screening, equivocal articles were screened by at least two team members.

Data Extraction

We extracted bibliographic data, eligibility criteria, enrollment years, study duration, and sample size for all studies. For Key Question 1, we extracted data that allowed reconstruction of trends over time in incidence and mortality, as well as patient-, tumor-, and system-level characteristics of interest. We extracted data into tables of 5-year bins (e.g., 1980-84, 1985-89) from 1980 to 2010. We extracted reported statistical data regarding changes over time in factors of interest. For Key Question 2, we extracted data on patient- and tumor-level characteristics used as eligibility criteria, followup or monitoring parameters, and specific triggers for definitive treatment. We also extracted definitions of disease progression. For quantitative studies (multivariable models) related to Key Question 3, we extracted the definition of the observational strategy, factors of interest, and effect sizes. For qualitative studies (surveys) related to Key Question 3, we extracted the specific survey approach used, the definition of the observational strategy addressed, the qualitative summary of the key study findings, and information to assess the study validity (e.g., survey response rate, survey validation). For Key Question 4, we extracted details about the study population (including eligibility criteria and baseline characteristics), specific interventions compared, outcome definitions, study design, and effect sizes of outcomes of interest.

Quality Assessment

We formally assessed methodological quality only for studies included for Key Question 4. Studies were graded using standard AHRQ EPC methodology with a three-level grading system (A, B, or C). For RCTs, we primarily considered the methods used for randomization, allocation concealment, and blinding, as well as the use of intention-to-treat analysis, the report of dropout rate, and the extent to which valid primary outcomes were described and clearly reported. Only RCTs and prospective comparative studies could receive an A grade. Retrospective studies could be graded either B or C. For all studies, we used the following in our assessment (as applicable): the report of eligibility criteria, the similarity of the comparative groups in terms of baseline characteristics and prognostic factors, the report of intention-to-treat analysis, important differential loss to followup between the comparative groups or overall high loss to followup, and the validity and adequacy of the description of outcomes and results. Quality A studies are those judged to have the least likelihood of bias and are considered the most internally valid. Quality C studies have a substantial risk of bias and may not be valid. Quality assessment was performed by the team member responsible for primary data extraction. The quality grade was confirmed by at least one other team member.

Data Synthesis

All included study data were tabulated into summary tables (provided in the report appendixes) that succinctly describe the important study characteristics and their findings. Time-trend data for Key Question 1 were graphed over the interval of interest (1980–2010). Although

we considered generating forest plots for comparative effectiveness data for Key Question 4, the data were inadequate for forest plots to be informative (i.e., there were generally only one or two studies addressing a specific question).

Grading the Body of Evidence

We graded the body of evidence only for the comparative effectiveness review portion of the systematic review (i.e., Key Question 4). We used standard AHRQ EPC methodology. We assessed the risk of bias of the studies based on their study design and methodological quality, the consistency of data across studies, the applicability of the studies to the U.S. population of men with localized prostate cancer, potential problems with measurement of outcomes in studies, and the precision and sparseness of data. The strength of evidence was rated on a four-level scale: High, Moderate, Low, and Insufficient. Ratings were assigned based on our level of confidence that the evidence reflected the true effect for the major comparisons of interest.

Results

Key Question 1. How have the patient population and the natural history of prostate cancer diagnosed in the United States changed in the last 30 years?

We identified 79 relevant primary observational studies and one systematic review. Of the primary observational studies, 51 analyzed the SEER database or a subset of its component registries, 9 the linked SEER-Medicare database, 11 the Cancer of the Prostate Urologic Research Endeavor (CaPSURE) database, 5 the National Cancer Database (NCDB), and 3 examined other large U.S.-based databases. In addition, we queried the online SEER database and reviewed the latest version of the Cancer Statistics report prepared annually by the American Cancer Society,[1] and a recent SEER Survival Monograph.[2]

Trends in Prostate Cancer Incidence

Prostate cancer incidence rates rose between 1975 and 1992 (from approximately 100 to more than 240 new cases per 100,000 men per year), and then fell until around 1995. After a period of nonsignificant increase from 1995 to 2000, rates declined again from 2000 to 2007 (to the current level of approximately 156 new cases per 100,000 men per year).[3] Overall, 33 studies provided information on trends of prostate cancer incidence stratified by factors relevant to Key Question 1.

Age

Eleven studies (covering 1969-2005) reported prostate cancer incidence rates according to age group. Collectively, they indicated an increase within all age groups until 1992-93 and then a decline until 1995-99. One study reported the following: compared to the pre-PSA era (1986), the incidence rates in 2005 were 3.64 times higher for men aged 50-59 years, 1.91 times higher for men aged 60-69, and 1.09 times higher for men aged 70–79 years, but only 0.56 times as common in men 80 years or older.

Race/Ethnicity

According to the 17 studies (covering 1973–2005) reporting incidence rates stratified by race/ethnicity, all groups experienced increases in prostate cancer incidence since the mid-1980s. The incidence rate appears to have peaked in the early 1990s for all racial/ethnic groups.

Tumor Stage

Fifteen studies (covering 1969–2005) reported incidence data, stratified by tumor stage. These studies consistently demonstrated that early-stage (localized and regional) prostate cancer cases were responsible for the observed increase in prostate cancer incidence from the mid-1980s up to the mid-1990s. Studies also consistently indicated a decrease in incidence rates for all disease stages from the mid-1990s to 2000. A single study investigated trends in incidence stratified by T stage, and demonstrated that compared to 1988–89 the incidence rate in 2004–05 reflected an increase of 76 cases per 100,000 person-years for T1 tumors (from a baseline rate of 42.3 cases per 100,000 person-years) and 11.2 cases per 100,000 person-years for T2 tumors (from a baseline of 95 cases per 100,000 person-years). In contrast, the incidence of T3 or T4 tumors (combined) decreased by 47.1 cases per 100,000 person-years (from a baseline of 55.5 cases per 100,000 person-years) over the same time period.

Tumor Grade

Five studies (covering 1973–2005) reported prostate cancer incidence rates stratified by tumor grade (level of tumor differentiation or Gleason score). In studies using SEER data, the increase in prostate cancer incidence observed from the mid-1980s to early 1990s was mainly due to an increase in the incidence rate of moderately differentiated tumors (corresponding to Gleason score 5–6). A single study using SEER data after 2000 reported a continued increase in incidence rate of moderately differentiated tumors from 1988 to 2005 and a concomitant decrease in the incidence rate of well differentiated tumors (corresponding to Gleason score 2–4).

Trends in Prostate Cancer Mortality and Survival Rates

For the overall U.S. population, the NCI's Cancer Trends Progress Report (2009/10) indicates that, after increasing from 1975 to 1991, prostate cancer death rates fell from 1994 to 2007. The baseline prostate cancer mortality in 1975 was 31 deaths per 100,000 men per year and has declined to the current level of approximately 24 deaths per 100,000 men per year. Overall, 17 studies provided information on trends of prostate cancer mortality or survival after diagnosis.[3]

Age

Nine studies (covering 1969–2003) reported prostate cancer mortality and survival rates stratified by age group. Collectively, they demonstrated decreases in the mortality rate for all age groups between the early 1990s and 1999. One study of prostate cancer-specific survival indicated that over time (1988–95) the proportion of patients diagnosed with prostate cancer who died of their cancer had decreased (i.e., patients with prostate cancer have increasingly died of other causes) across all the age groups considered (>50 years old).

Comorbidity (Other Primary Cancers)

One study demonstrated that throughout 1988 to 1995 prostate cancer patients with other primary tumors were consistently less likely to die of prostate cancer compared to patients with no other primary tumors.

Race/Ethnicity

According to 15 studies (covering 1969–2000) reporting prostate cancer mortality and survival rates stratified by race/ethnicity, mortality rates among blacks were consistently higher compared to those of non-Hispanic whites. Similarly, black patients with prostate cancer were at higher risk for death due to prostate cancer compared to non-Hispanic whites, although the difference between these two groups appeared to decrease over time.

Tumor Stage

Six studies (covering 1969–2001) reported information on prostate cancer mortality stratified by tumor stage. Data indicated that over time the proportion of deaths due to prostate cancer among patients diagnosed with the disease had decreased, particularly for patients with early-stage (localized or regional) disease at diagnosis.

Tumor Grade

Three studies reported information on prostate cancer mortality stratified by tumor grade (of patients diagnosed in 1973–97). One study demonstrated that the probability of dying from prostate cancer among patients diagnosed with prostate cancer decreased during the study period (1988–95). Although the decrease was observed for all cancer grades, it was more pronounced among patients with well- and moderately-differentiated tumors. The second study demonstrated that, compared to patients with well differentiated tumors, patients with moderately and poorly differentiated tumors had a higher probability of prostate cancer death (more than 2-fold and more than 4-fold higher, respectively). These differences were relatively constant over the time period covered by the study (1988–95). The third study demonstrated that patients with low grade tumors have the highest relative survival compared to those with higher grade disease; improvements in relative survival over time were observed for all tumor grades.

Patient, Tumor and System-Level Characteristics at Diagnosis

We identified 52 observational studies reporting on patient characteristics at presentation.

Age

Twenty-one studies (covering 1973–2005) reported information on patients' age at presentation. Among seven studies evaluating average age at diagnosis of prostate cancer, five found reductions in the average age of patients whereas two studies did not report any changes during their respective time periods. Fifteen studies that evaluated the distribution of patients' ages into discrete categories generally supported a trend toward younger age at diagnosis (the effect was significant in four of the six studies reporting statistical tests).

Comorbidity

Two studies (covering 1997–2003) reported information on comorbidities at diagnosis. The CaPSURE analysis found no statistically significant difference in the distribution of patients with no, one or two, or three or more comorbidities, when comparing 1997–99 versus 2000–03. The

POCS analysis found that the proportion of patients with no comorbidity increased from 78.3 percent in 1998 to 87.4 percent in 2002.

Race/Ethnicity

Eighteen studies (covering 1973–2003) reported information on race/ethnicity. No consistent pattern in the racial or ethnic distribution of cases over time was found: some studies indicated that the number of whites increased over time, others that it remained stable, and others that it decreased. Studies using the same database often provided discrepant results even for overlapping time periods; these findings may be a reflection of the different inclusion criteria used in each study.

Tumor Stage

Twenty-two studies (covering 1973–2007) reported information on trends in the distribution of prostate cancer stage at diagnosis. Studies reporting on cancer stage consistently demonstrated decreases in the proportion of patients presenting distant disease and concomitant increases in the proportion of patients with localized or regional disease, over their respective time periods. Studies consistently demonstrated reductions in the proportion of patients presenting with higher T stages. The two studies reporting on T1/T2 tumors both demonstrated a decrease of T1a/T1b tumors and T2a tumors and an increase in T1c tumors.

Tumor Volume

We did not find data on changes in tumor volume.

Tumor Grade

Sixteen studies (covering 1973–2006) reported information on trends in the distribution of tumor stage at diagnosis. These data consistently demonstrated reductions in the proportion of patients diagnosed with well- or poorly-differentiated tumors (including undifferentiated tumors) with concomitant increases in the proportion of patients with moderately-differentiated disease.

Prostate Specific Antigen

Eight studies (covering 1989–2007) reported PSA data. These studies indicated that PSA values at diagnosis have decreased over time (i.e., that a larger number of patients are currently diagnosed with PSA concentrations below 10 ng/mL).

Biopsy Frequency

Four studies (covering 1982–2001) reported information on trends in the performance of prostate biopsies. A study based on the SEER-Detroit database reported that the proportion of prostate cancer patients diagnosed through biopsy (compared to those diagnosed through other procedures, such as transurethral resection of the prostate) increased over time. A similar trend was observed in a study based on the SEER-New Mexico registry. A SEER-Medicare study demonstrated an increase in the age-adjusted rate of biopsy procedures (from 685 to 2600 per 100,000 men) between 1986 and 1991. An update from the same database reported that there was no statistically significant change in the population biopsy rate between 1993 and 2001.

Number of Cores

One study examined trends in the number of biopsy cores obtained during diagnostic workup, and found that between 1997 and 2002 the average number of cores obtained per patients had increased by 0.41 cores annually (from a mean of 7.5 to a mean of 9.8 cores per patient).

Histopathologic Grading Changes

One study reported the results of regrading in 2002-04 pathology slides from patients diagnosed in 1990-92. The regrading resulted in the assignment of significantly higher Gleason scores compared to the original readings (mean score increase from 5.95 to 6.8).

Differences in Geographical Access and Other System-Level Factors

Four studies (covering 1986–2003) reported information on changes in the distribution of patients by system-level factors. Among three studies on trends in the distribution of patients' insurance status at diagnosis, the two CaPSURE analyses demonstrated a decrease in the proportion of patients with Medicare coverage at the time of diagnosis over the time periods covered (1997–2003 and 1989–2001). The POCS analysis did not demonstrate a change in the distribution of insurance status over time (1998–2002). An analysis of POCS comparing 1998 to 2002 reported an increase over time in the number of patients residing in areas of higher median income. Patterns in the distribution of income are difficult to interpret because sampling strategies changed and different regions were included at the different time points. An analysis of NCDB found little evidence of change in the distribution of patients by hospital caseload over time (1986–87 and 1992).

Trends in Treatment Patterns

Among the 21 studies (covering 1973–2008) from which data could be gleaned regarding treatment patterns over time, most demonstrated decreasing trends in the proportion of patients being managed with observational strategies of no active treatment (AS, WW or expectant management), with or without androgen deprivation therapy (ADT). In all seven studies providing data since 2000, the proportion of patients receiving AS or WW was less than 10 percent; this also held true for the subgroups of patients with "low-risk disease" investigated in two studies.

Key Question 2. How are active surveillance and other observational management strategies defined?

The terms AS and WW (as well as others) have been used by investigators to denote strategies both with and without curative intents. There is a broad spectrum of approaches for observational strategies described in research publications. For the purpose of operationalizing the process of summarizing the various definitions, we divided protocols into those clearly described as having curative intent and those in which their aims were either unclear or primarily palliative, regardless of how these regimens were labeled. This categorization was applied for practical reasons, not to suggest what the definitions or protocols for AS, WW, or any other observational strategy should be.

Strategies With Curative Intent

Sixteen unique cohorts reported criteria and protocols for AS (i.e., studies that met our criteria of monitoring triggers for curative treatment of prostate cancer other than symptom progression). In all cohorts, AS was offered to men with low-risk or clinically localized prostate cancer.

Eligibility Criteria

Other than restriction to men with clinically localized prostate cancer (T1 or T2), eligibility criteria for AS varied across the cohorts. The most commonly used eligibility parameters were Gleason score (12 cohorts), PSA (10 cohorts), and number of biopsy cores positive for cancer (8 cohorts).

Age

Only three studies used age as an eligibility criterion, restricting subjects to men under age 70 or 80 years.

Gleason Score

Twelve cohorts based eligibility for AS on the Gleason score. Generally, cohorts used Gleason score of 6 or less (no pattern 4 or 5). Three cohorts allowed Gleason pattern 4.

Number of Cores Positive for Cancer

Eight cohorts used a maximal number of biopsy cores positive for cancer as part of the eligibility criteria for AS. Five cohorts allowed two or fewer cancer-positive cores; three cohorts allowed three or fewer. Some cohorts used sextant, some octant, and some extended (>10 cores) biopsies.

Percentage Cancer Involvement in Each Core

Five cohorts used "low-volume disease" as part of the patient eligibility criteria for AS. In three cohorts, the definition of "low-volume disease" was involvement of less than half of any individual core with cancer. In the other two cohorts, the criterion percent of biopsy cores with cancer involvement was described variably as less than half of two biopsy cores, less than 20 percent in one or two biopsy cores, and less than 33 percent of biopsy cores.

Prostate-Specific Antigen

Ten of 16 cohorts used PSA as part of the eligibility criteria for AS. Three PSA thresholds were used: ≤10 (7 cohorts), ≤15 (3 cohorts), and ≤20 (2 cohorts) ng/mL. Two cohorts used PSA density (PSA per volume of prostate tissue) thresholds.

Imaging

One cohort required a chest radiograph as part of patient eligibility criteria, and another cohort noted that magnetic resonance imaging was selectively used at diagnosis.

Behavioral Indicators

No behavioral indicator was used explicitly as a criterion for AS enrollment.

Followup Protocols

All 16 cohorts included regular PSA testing in the followup protocol. Twelve cohorts included regular digital rectal examination (DRE). Fourteen cohorts performed routine rebiopsy. The testing frequency of PSA, DRE, and rebiopsy varied across the cohorts. One cohort also incorporated a regular bone scan schedule. Criteria for recommending curative treatments varied across the cohorts. The recommended treatments were not standardized and were left at the discretion of treating physicians and patients in many of the cohorts.

Gleason Score

Twelve cohorts described using the Gleason score as part of their monitoring criteria for disease progression. Generally, disease progression was defined as a Gleason score or pattern greater than those used in the eligibility criteria for AS.

Number of Cores Positive for Cancer

Eight cohorts included the minimum number of biopsy cores positive for cancer as part of their monitoring criteria for disease progression. Two criteria were used: three or more and greater than four positive biopsy cores (6 and 3 cohorts, respectively). Rebiopsy frequencies varied across the cohorts.

Percentage Cancer Involvement in Each Core

Six cohorts used more than 50 percent cancer involvement in each biopsy core as part of monitoring criteria for disease progression. Two other cohorts considered an increase in tumor volume as part of the monitoring criteria for disease progression, but specific percentage cancer involvement was not reported.

Prostate-Specific Antigen

All 16 cohorts included regular PSA testing in the followup protocol. Six cohorts considered rising PSA and/or PSA kinetics as part of triggers for treatment but did not specify the detailed criteria. Nine cohorts used a variety of PSA triggers for treatment.

Imaging

One cohort performed an annual bone scan for the first 2 years and biennially thereafter. Another cohort reported that magnetic resonance imaging of the prostate was selectively performed every 1 to 3 years during followup.

Behavioral Indicators

No study used a formal assessment of any behavioral indicator for triggering active treatment as part of their followup protocol, but one cohort reported that some patients requested treatment due to anxiety related to increasing PSA concentration.

Observational Management Strategies With Palliative Intent

Thirteen cohorts reported followup protocols for patients who initially received no treatment and who were subsequently treated only for symptomatic progression.

Eligibility Criteria

The six cohorts that enrolled patients in the pre-PSA screening era primarily based enrollment on clinical staging alone. In the PSA era, the seven cohorts mostly enrolled patients with stage T1 or T2 cancer or without evidence of nodes or metastases. The commonly used patient eligibility criteria were PSA (5 cohorts), age (4 cohorts), Gleason score (4 cohorts), and normal bone scan findings (4 cohorts).

Age

Four cohorts included age as part of their eligibility criteria. The different thresholds used were less than 75 years (2 cohorts), less than 85 years, and between 50 and 75 years.

Gleason Score

Four cohorts used Gleason score thresholds. Three used a threshold of less than 8. One required that less than 25 percent of the tumor was Gleason grade 4 and less than 5 percent grade 5.

Number of Cores Positive for Cancer

No cohort used this factor.

Percentage Cancer Involvement in Each Core

No cohort used this factor.

Prostate-Specific Antigen

Five cohorts used PSA as part of their eligibility criteria, with thresholds of less than 50 ng/mL (4 cohorts) and less than or equal to 15 ng/mL (1 cohort).

Imaging

Four cohorts required normal bone scan findings. One of these cohorts also required normal chest radiograph findings.

Behavioral Indications

No cohort used this factor.

Followup Protocols

Five of the six cohorts in the pre-PSA screening era included regular prostate acid phosphatase (PAP) testing and bone scan in the followup protocol. The sixth cohort reported regular PSA and DRE in the followup protocol for patients who received no treatment after the introduction of PSA in 1990. All seven cohorts in the PSA screening era included regular PSA testing. Compared with AS cohorts (see previous section), rebiopsy was not commonly included in the followup protocol among WW cohorts.

Gleason Score

No cohort used this factor.

Number of Cores Positive for Cancer

No cohort used this factor.

Percentage Cancer Involvement in Each Core

No cohort used this factor.

Prostate-Specific Antigen

Three cohorts formed in the pre-PSA screening era reported that PSA testing became part of followup protocol after PSA became available. All six cohorts in the PSA screening era included regular PSA testing as part of followup protocol. However, rising PSA concentration alone was not used as a trigger for treatment in five cohorts. The sixth cohort reported that "hormonal manipulation was demanded by the protocol when the PSA rose to 50 ng/mL."

Imaging

Five cohorts in the pre-PSA screening era included regular bone scan in the followup protocol. One cohort also included regular chest and skeletal radiographs in the followup protocol. Another cohort reported that computed tomography of the pelvis was conducted infrequently. Three cohorts in the PSA screening era included regular bone scans and chest radiographs in the followup protocol. Another cohort reported that all patients underwent "multiple bone scans" during followup.

Behavioral Indications

No cohort explicitly used this factor.

Implicit in the Key Question is a comparison between AS and other observational strategies in the modern PSA era. Thus, we compared the 16 unique cohorts reporting formal protocols to monitor triggers for curative treatment with the 7 unique cohorts of other observational strategies with primarily palliative intent in the PSA screening era. Enrollment into AS protocols more commonly used Gleason score as a threshold than other observational strategies. They also used the number and percentage of cores positive for cancer as a threshold, while none of the other strategies used these factors. Both sets of strategies generally used some sort of PSA criteria, but the thresholds in AS were generally lower (10-15 ng/mL) than the other observational strategies (15 or 50 ng/mL). AS protocols had more clearly defined followup processes than other observational management strategies, with explicit indications for curative treatment including increase in Gleason scores, number and percentage of positive cores (on rebiopsy), and PSA velocity. AS protocols generally did not include imaging as part of their followup processes. In contrast, other observational strategies typically included imaging in their followup, specifically bone scan and chest radiography. They also generally did not employ rebiopsy but did use PSA in their followup. Comparison of the followup frequencies between AS and other observational strategies showed that PSA testing and DRE were common in both strategies, but somewhat more frequent with AS protocols, at least within the first year of followup.

Key Question 3. What factors affect the offer of, acceptance of, and adherence to active surveillance?

We included three types of studies to address this Key Question. We included multivariable database analyses of predictors for the offer or acceptance of or adherence to AS (or WW). We included survey or questionnaire studies addressing the same issues. We also searched for experimental studies evaluating the effect of tools, such as decision aids, on the offer or acceptance of or adherence to AS (however, no such studies were found). Of note, the outcomes

of many of the examined studies were either treatment with an observational strategy or interruption (cessation) of the observational strategy. Studies generally did not directly analyze the offer or acceptance of or adherence to AS.

Primary Care

One survey of New Zealand general practitioners found that 45 percent would offer observational management if the patient's life expectancy was <10 years, but only 3 percent would offer observational management to patients with a longer life expectancy. Five surveys of patients reported that their physician's treatment recommendation was the most influential factor in deciding on their treatment. In one survey, 81 percent of men on observational management who ultimately received active treatment believed that the treatment was favored by their physicians; in contrast, only 24 percent of the physicians' notes documented that the physician recommended treatment.

Diagnosing Physician

One survey of patients on observational management strategies reported that observational management strategies were offered by 36 percent of the physicians who had made the initial diagnosis.

Consultant—Second Opinion

One survey was of men diagnosed with early-stage cancer. They had not yet decided on treatment and were recommended by their urologists to seek a second opinion. None of the men followed through with the recommendation to seek a second opinion, but the offer reinforced their trust and confidence in their urologists. A survey of Australian men who had a urological consultation reported that 71 percent of the urologists discussed observational management strategies, compared with 92 percent who discussed RP and 87 percent RT. One survey of urologists regarding men with localized cancer and few comorbidities found that 4 percent preferred observational management strategies; two-thirds preferred RP. The same study reported that 20 percent of patients thought that treatment options were not discussed, while only 1 percent of the urologists thought so. In a survey of men and their urologists, the urologists, in an initial consultation setting, recommended observational management strategies to 25 percent of men and offered 0.5 more treatment options than the urologists in a second opinion visit setting, who recommended observational management strategies to only 16 percent of men.

Clinical Factors

One survey of urologists and radiation oncologists reported that about 10 to 20 percent would recommend observational management strategies for a 65 year old man with a low PSA, a Gleason score of 4 or 5, in good health, with negative DRE, and no evidence of nonlocalized disease. Almost none would recommend observational management strategies for those with higher PSA or Gleason scores. The responses of urologists and radiation oncologists did not differ significantly. Numerous multivariable analyses found that receipt of observational management strategies was predicted by older age, an increased number of comorbidities, lower Gleason score, well-differentiated tumor, lower stage disease, lower PSA, or low-risk on the D'Amico scale. Multivariable analyses also found that interruption of observational management strategies was predicted by higher stage disease, higher PSA at diagnosis, decreased free-to-total

PSA ratio, or more rapid PSA increase, but not comorbidities or Gleason score; two of four studies found an association with younger age and one of three with higher D'Amico risk score.

Family Involvement

In two surveys, advice from family and friends was the most influential factor in deciding treatment in 19 and 9 percent of men surveyed. In a focus group, half the men reported relying on influential others to make a treatment decision (either for or against observational management strategies). In an open-ended interview of men with localized disease, 4 percent reported that family opinions were a reason for not choosing observational management strategies.

Personal Preferences

One analysis compared men who refused randomization, but instead selected AS (i.e., men who did not participate in the trial), to men in the trial who were randomized to AS. It found that lower baseline anxiety was associated with the decision to choose AS (and not be randomized). Four surveys found that concern for treatment side effects (primarily impotence and incontinence) were reasons that men chose observational management strategies. Three multivariable analyses found predictors of choosing observational management strategies included the desire to avoid side effects or having current bowel problems, urinary dysfunction, or other urinary conditions; sexual dysfunction was predictive of choosing RT over observational management strategies. One multivariable analysis also reported that increased anxiety was associated with an increased probability of interruption of observational management strategies.

Risk Perceptions

One set of interviews in men with low-risk prostate cancer reported that physician description of prostate cancer affected treatment choice. One survey of men with early stage prostate cancer reported that men who chose RP over RT or observational management strategies perceived prostate cancer as a significantly more serious disease. Another survey of men with localized prostate cancer reported that fear of consequences was the most common reason for not selecting observational management strategies.

Family History

Two multivariable analyses reported that family history was not a significant factor in predicting interruption of observational management strategies.

Social Support

Four multivariable analyses reported that not being married or in a permanent relationship were associated with an increased probability of receiving observational management strategies. One survey of couples in which the men were diagnosed with early-stage cancer but had not yet decided on treatment concluded that couples ruled out options based on both formal (provided by the physicians) and informal (provided by family and friends) information, and that they also "considered both their own individual histories and concerns and their shared life experiences." One multivariable analysis reported that marital status was not associated with time to interruption of observational management strategies.

Insurance Type

Two multivariable analyses reported that having Medicare insurance increased the probability of receiving WW/AS compared with private or Veterans Administration insurance. One analysis reported that having preferred provider organization or health maintenance organization coverage decreased the probability of receiving observational management strategies versus RP. It also reported that Medicare supplemented with fee-for-service, health maintenance organization, or preferred provider organization coverage decreased the probability of receiving observational management strategies versus RP. One multivariable analysis reported that insurance status was not a significant factor in predicting interruption of observational management strategies.

Availability of Technology

No study addressed this factor.

Small Area Variation

No study addressed this factor.

Regional Variation

One multivariable analysis comparing the registries in the National Cancer Institute's Patterns of Care study claimed that men who resided in New Jersey had an increased probability of receiving observational management strategies compared with men in California (excluding three major cities). Comparisons among other registries were nonsignificant. Another multivariable analysis reported that men in Northeast had a decreased probability of selecting observational management strategies (versus active treatments) compared with men in California (excluding three major cities).

Urban Versus Rural

One survey of men with prostate cancer in North Carolina reported that there was no significant difference between urban and rural residents in North Carolina as to whether the option of observational management strategies was discussed with their physicians. One multivariable analysis reported that men who resided in urban areas (vs. rural areas) had a decreased probability of receiving observational management strategies versus RP or RT. The survey in North Carolina reported that there was a difference in whether physician recommendation was the most influential factor in the treatment decision between urban and rural residents (62 percent vs. 44 percent, respectively).

Academic Centers Versus Private Practice

One multivariable analysis reported that treatment facility status (academic vs. community practice) was not a significant factor in predicting receiving observational management strategies versus active treatment.

Risk Assessment, Predictive Models

No study addressed this factor.

Decisionmaking Tools and Aids Specifically for AS

No study addressed this factor.

Key Question 4. What are the comparative short- and long-term outcomes of active surveillance versus immediate treatment with curative intent for localized prostate cancer?

In order to understand the effectiveness of AS relative to active treatment options, studies of AS need a control group for comparison. As such, we did not include single-arm AS cohort studies, which cannot address comparative effectiveness questions. However, no study reported clinical outcomes specifically for AS management strategies with deferred treatment with curative intent versus immediate definitive treatment. Therefore, there is insufficient evidence to evaluate the comparative effectiveness of AS management with curative intent versus immediate definitive treatment in men with localized prostate cancer.

Faced with a lack of studies comparing AS to immediate active treatment, we elected to evaluate studies that compared other observational management strategies (largely resembling WW) with immediate treatment. In addition to previously published systematic reviews and evidence reports, our searches identified two updates from multicenter RCTs (four publications, three on clinical/quality of life outcomes and one on costs) and 16 cohort studies (3 prospective and 13 retrospective). Notably, the majority of evidence for this Key Question came from retrospective analyses of observational studies. Confounding by indication is likely in these studies, due to the differences in patient characteristics and risk profile between patients treated with observational strategies and those who received active treatment.

Observational Management Strategies Versus Radical Prostatectomy

Studies generally reported that men treated with RP had lower all-cause or prostate cancer-specific mortality rates than men on WW. The development of metastatic disease was assessed by a single study that found a significant benefit for RP compared to WW. Morbidity of primary treatment was reported by two studies that suggested an increased risk for urethral stricture (and procedures to treat it) were less likely among patients on observational management. Quality of life (QoL) was reported in three studies; the results varied across different domains of QoL measure.

Observational Management Strategies Versus Radiation Therapy

Studies generally reported that men treated with RT had lower all-cause mortality rates than men on WW. One study reported prostate cancer-specific mortality information and did not find a statistically significant difference between RT and observational management. No study reported on treatment comparisons for the development of metastatic disease. One study did not find a significant difference in morbidity between observational management and BT or EBRT. QoL measures and satisfaction with treatment were reported in four studies; the results varied across different domains of QoL measure.

Observational Management Strategies Versus Combined Active Treatments or Combined Radiation Treatment Modalities

One study reported that active treatments (RP, RT, and BT considered together) resulted in lower all-cause and prostate-cancer-specific mortality rates compared to WW. Morbidity of primary treatment was reported by only one study, which found that a group of patients receiving EBRT and BT (combination therapy) had a higher rate of receiving treatments for urethral stricture compared to a group managed using observational management strategies.

Costs

Short- and long-term costs appear to be higher for active treatment strategies (RP or RT) compared to WW; however, evidence originated from small studies (or studies where the subgroup of patients receiving observational management was small) using heterogeneous measurement methods. We did not identify any primary study comparing actual costs of AS versus active treatment strategies; economic modeling using U.S. prices suggested that within 10 to 15 years of diagnosis AS may be less costly compared to active treatments; a study using a lifetime horizon indicated that AS may be associated with higher costs compared to RP and BT, but lower costs compared to intensity modulated RT (IMRT) and proton beam RT. We note that model based costs are sensitive to the model assumptions and choice of inputs.

Key Question 5. What are the research needs regarding active surveillance (or watchful waiting) in localized prostate cancer?

The evidence directly addressing the four principal Key Questions is largely incomplete. There is not yet consistency among clinicians or researchers as to the definitions of AS or WW, the standard protocols for the interventions, or how to manage patients whose cancers show signs of progression. There are also many gaps in the evidence regarding the numerous specific factors and subgroups of interest to the conference.

This review implicitly assumes that it is possible to identify men who are at sufficiently low risk of progression of their prostate cancer that AS can be a safe and appropriate option for them. However, additional basic and clinical research is needed to more accurately classify or predict those men whose diseases are indeed at a low risk of progression. These are the men that presumably would be most appropriate to consider offering AS.

Key Question 1—Patient Population and Natural History Changes in the Last 30 Years

Better understanding of time trends can be gained by improving the data collected and expanding the scope of major U.S. databases. In particular, stage and grade information are often incomplete, requiring researchers to create broad categories that place major limitations on analyses. The SEER database often appears inadequate for analyses on races other than blacks and whites; this may require adding new registries to SEER that better represent other races.

A misclassification bias is likely in the analyses of SEER using the “best available information” on staging information, because the “best available staging information” depends on the treatment the patients receive. Patients having surgery are staged more accurately than those with clinical or imaging staging alone. This bias could be reduced if the SEER database maintained the staging information available prior to surgery.

Key Question 2—Definition of Active Surveillance

Little new research is needed to address how active surveillance has been defined by researchers. However, interpretation of future studies would be best served if there were a standard, agreed-upon definition of AS that clearly distinguished it from WW and other forms of withheld or noncurative treatments. A consensus conference may be the most appropriate forum to define AS. Features of the definition will need to include (1) the goal or intent of the intervention; (2) the "eligibility criteria," a determination of which patients should be offered AS based on disease and patient characteristics; (3) the "followup protocol," the minimum set of tests that should be followed and their timing; and (4) the criteria or triggers for stopping AS to seek definitive treatments.

Assuming that AS is an intervention plan that many patients may select (if offered) in order to avoid the side effects from immediate invasive treatment for a potentially nonlethal disease, it would be desirable to determine the best AS protocol that would minimize prostate-cancer specific morbidity and mortality, and that patients and caregivers would adhere to. This best AS protocol should be investigated by randomized or other prospective comparative studies that directly compare different protocols. Examples of comparisons for future trials could include use of different combinations of followup testing, different timing for the tests, and different definitions of progression that would determine when curative treatment is offered. The outcomes of greatest clinical importance are those that are most pertinent to patients' health, well-being, and longevity. Examples include all-cause mortality, prostate-cancer-specific mortality, symptomatic disease, urological and other complications (from testing, treatment, or deferring treatment), quality of life, anxiety, and family dynamics. Also of interest would be overall costs, use of resources, and numbers of negative invasive tests (i.e., biopsies showing no progression, thus arguing they were unnecessary).

At a minimum, future study reports should be very explicit and clear as to their definitions of AS (or WW), the goals of the intervention, the exact protocols, the exact definitions of progression, how and when protocols or standards changed during their study (and why), and how often and why patients and clinicians chose not to follow the protocols.

Key Question 3—Factors That Affect Offer of, Acceptance of, and Adherence to Active Surveillance

Current databases tend to have data only about what treatment patients received and when. Therefore, whether different treatment options were offered to them, whether they accepted those options, and whether they adhered to their initial choices could only be inferred. Even the best analysis of predictors of initial treatment cannot adequately address this Key Question. Thus, full statistical analyses of predictors will require the prospective collection of data specifically about what interventions were offered to each patient, which treatments the patients accepted, and when they chose to receive curative treatment despite lack of evidence of progression. These datasets will need to be sufficiently large to allow for testing of multiple predictor variables. In addition, future studies should only perform complete analyses of all treatment options without arbitrarily grouping treatments or selectively excluding treatments. This will minimize bias and increase clarity about what is being tested.

Future database analyses and prospective observational studies should focus on those predictors that are amenable to change or that can be acted upon. Researchers should avoid

interpreting analyses to suggest that men with certain demographic (or other nonmodifiable) features are more likely to accept treatment and thus other men should not be offered treatment.

Further surveys of patients, their families, and their clinicians are warranted. To improve reliability, these should be adequately powered to ensure that sufficient numbers of men were treated with different interventions and to allow full analyses of the tested predictors. Studies should use established methods including standardized qualitative research designs and, ideally, validated questionnaires to elicit preferences. Studies of this sort also need to consider the overall adequacy of discussion with patients regarding management options—and documentation of those discussions. Adequate documentation of these discussions will surely improve the veracity of some of these survey data.

Future Key Questions of interest could include comparisons of interventions that improve the likelihood that eligible men are offered AS, that improve acceptance of AS, and that improve adherence with AS, so long as it remains the most appropriate treatment. Arguably, it is more important to first establish how to successfully get men offered, accepting, and adhering to AS, before determining which men are at greatest risk of failing to receive AS.

Key Question 4—Active Surveillance Versus Immediate Curative Treatment

A randomized controlled trial with long followup would provide the best evidence to adequately assess the differential effects between AS and immediate curative treatment. The least biased, most reliable study design comparing two interventions is the randomized controlled trial that adheres to modern methodological standards. Outcome assessors—particularly those who conduct psychometric testing—should be blinded. The primary outcomes of interest should be patient-centered outcomes, including mortality and disease-free survival, psychometric measurements, adverse events, resource utilization, and costs. Trials need to be of sufficiently long duration to collect data on the clinically relevant outcomes. However, limited resources and possible difficulties recruiting patients may preclude such a study.

In lieu of randomized trials, adequate findings may possibly be extracted from long-term databases with prospectively collected data. However, these studies, too, should use AS protocols that are defined a priori and undergo minimal change over time or between centers. The determination of which patients are potentially eligible for AS should also be made a priori. These studies will need to use multivariable analyses, propensity scores, or other validated methods to adjust for the broad range of factors that affect the decision to use AS. We do not believe that retrospective studies are capable of providing adequate data for unbiased analyses.

Subgroup analyses of both randomized trials and prospective observational comparative studies should be conducted to look for particular sets of men who may benefit most (or least) from one approach or the other. Preferably, these subgroups should be considered a priori. The factors listed in Key Questions 1 and 2 form a good starting point to consider which subgroups may be of interest. In addition, future studies that could uncover better bio- and imaging markers of indolent versus aggressive disease, and thus could better stage patients as having either low or high risk disease, are necessary to better inform which patients are most likely to benefit from observational versus active treatment.

Discussion

Prostate cancer epidemiology is affected by population-level trends, such as the aging of the U.S. population, but also by changes in the application of screening and diagnostic technologies among the population at risk. Keeping these caveats in mind, studies indicate that men in all racial/ethnic groups experienced increases in prostate cancer incidence since the mid-1980s, with rates peaking in the early 1990s. For all groups, incidence rates declined between the early-1990s and 1999. Studies have consistently demonstrated that early-stage (localized and regional) prostate cancer cases were responsible for the observed increase in prostate cancer incidence from the mid-1980s up to the mid-1990s. Studies also demonstrated decreases in the prostate cancer-specific mortality rate for all age groups between the early-1990s and 1999. Mean age of diagnosis has also decreased over time for both blacks and whites. Another consistent trend in SEER data has been the decrease in low-grade (Gleason score 2–4) and high grade (≥7) tumors, and a concomitant increase in intermediate grade tumors (Gleason 5–6). It has been hypothesized that this effect is due to changes in histopathological grading guidelines, a preference towards avoiding assigning Gleason 2–4 scores based on prostate cancer biopsy samples, and PSA test's ability to detect moderately differentiated tumors with higher accuracy (compared to poorly-differentiated tumors). Most studies demonstrated decreasing trends in the proportion of patients being managed with strategies other than RP or RT throughout their respective time periods. Studies explicitly reporting on AS/WW-type strategies indicated decreases in the proportion of patients receiving such treatments over time; this was true even for subgroups of men with "low-risk disease."

There is not yet a consensus among clinicians or researchers as to the definitions of AS or of WW, the standard protocols for the interventions, or how to manage patients whose cancers show signs of progression. This is evidenced by the 16 unique cohorts formed in the PSA screening era that used different formal protocols to monitor triggers for curative treatment of prostate cancer. In all these cohorts, AS was offered to men with low-risk or clinically localized prostate cancer although no uniform criteria were used to identify these men, with the exception that no cohorts enrolled patients with tumors of a clinical stage greater than T2. They employed different combinations of periodic DRE, PSA testing, rebiopsy, and/or imaging findings to determine different thresholds used for seeking definitive treatments. The AS followup protocols also varied across these cohorts.

Owing to the variation in usage of the terms AS and WW, and their intended and often mixed treatment objectives (both curative and palliative), it is often difficult when reviewing the studies to know whether patients received true AS or WW, or were simply not treated (for a variety of reasons) or experienced delays in their treatment (and thus initially had no treatment).

Only two studies specifically examined factors related to men who were enrolled in an active monitoring protocol with triggers for curative treatments. The first found that the free to total PSA ratio and T stage were independent predictors of time to radical treatments in patients on the protocol, while initial PSA, PSA density, Gleason score, number of positive cores, and prostate volume were not independent predictors. The second study found that men with decreased baseline anxiety and higher socioeconomic status were associated with decreased probability of willingness to consent to randomization for AS versus definitive treatment (i.e., these men proactively selected AS). The rest of the heterogeneous studies reported on men who did not receive treatments or initial treatments. Therefore, whether they were on AS or WW could not be readily discerned. The following patient and clinical variables are potentially important in increasing the probability that a patient receives an observational management strategy: older

age, presence of comorbidities, lower Gleason score, lower tumor stage, lower diagnostic PSA, lower risk groups, or decreased baseline anxiety. The following patient and clinical variables are potentially important in increasing the probability that a patient interrupts an observational management strategy to seek definitive treatments: younger age, higher tumor stage, higher diagnostic PSA, higher PSA velocity, higher risk groups, or increased anxiety.

As most of these tentative conclusions are drawn from multivariable analyses of large databases that did not specifically address the factors that affect the offer or acceptance of or adherence to AS, whether different treatment options were offered to the patients, whether they accepted those options, and whether they adhered to their initial choices could only be inferred from whether they received the treatments or not. In addition, retrospective studies could not provide adequate data for unbiased analyses, because patient characteristics are strongly associated with initial treatment choice.

No trial provided results from comparisons of AS with RP or RT in men with localized diseases. One trial reported that men who underwent RP had lower mortality than men on WW; one trial reported that there was no difference in mortality comparing men having undergone RP with men in WW. Retrospective studies suggest that men on conservative management had a higher prostate-cancer-specific mortality than men treated with RP. Men who had RPs had more urinary complications than men on WW. Retrospective studies also reported that men treated with RT had lower mortality than men on WW. They also reported higher rates of urinary strictures in men treated with RT compared with men on WW. Definitive conclusions for men with low-risk disease on AS or WW versus RP or RT will have to await results from two ongoing trials: Prostate cancer Intervention Versus Observation Trial (PIVOT: Observation vs. RP) and Prostate Testing for Cancer and Treatment trial (ProtecT: AS vs. RP or RT).

Although cost calculations using retrospective primary data were performed using different methods and followup durations in each study, it appears that generally WW is associated with lower treatment costs compared with active treatment. On the other hand, model-based cost analyses of AS compared to active treatments suggested that AS costs accumulate over time. In these models, at 10 and 15 years of followup, AS appeared to be less expensive than active treatments. However; based on a model with a lifetime horizon, the costs of AS may exceed those of RP and BT with long term followup, and may be lower than those of IMRT or proton beam RT.

Introduction

In 2011, more than 240,000 men are projected to be diagnosed with prostate cancer and 33,000 to die from the disease in the United States.[1] In the United States, most cases of prostate cancer are detected via prostate-specific antigen (PSA) screening. The cancer is usually localized, and most tumors have low histological grades and low Gleason scores. Indeed, more than half of prostate cancers detected by PSA screening are expected to be early-stage, low-risk tumors.[4] Such cancers are an infrequent cause of death, and those affected are more likely to die of unrelated causes.

A number of immediate active treatment options are available for localized prostate cancer. Most commonly, radical prostatectomy (RP) or radiation therapy (RT), with or without androgen deprivation therapy (ADT) are offered with curative intent. However, the clinical benefit of immediate therapy with curative intent has not yet been demonstrated for localized prostate cancer in a PSA-screened population. It is likely that a large number of men are receiving treatment with curative intent without clinical benefit due to the slow progression of many prostate tumors.[4] Both surgical and radiation treatments result in significant short- and long-term adverse events, including impotence, urinary dysfunction, and other complications. Thus, determination of the appropriate management strategy for early-stage, low-risk prostate cancer is an important public health concern.

Active surveillance (AS) and watchful waiting (WW) are two observational followup strategies that forego immediate therapy in patients with prostate cancer. AS is curative in intent, while WW palliative. AS is appropriate in men with disease believed to be indolent and therefore may not require therapy. Because prediction tools are imperfect, these men are monitored closely and treated with curative intent at signs of progression or at patients' discretion. In this way, the considerable adverse effects of treatment are at best avoided and at least deferred. This approach is to be distinguished from men for whom treatment is deemed inappropriate due to comorbidity; for these men, WW is generally considered, as it offers the option of palliative therapy upon symptomatic disease progression. AS often entails a multifactorial followup of patients—monitoring of PSA values, digital rectal examinations (DRE), prostate imaging, and periodic prostate biopsies—while WW is commonly a relatively passive strategy, with interventions triggered by symptoms. However, there is a continuum of aggressiveness of followup for both AS and WW, as currently practiced. Even though the two terms are used commonly in the scientific literature, the underlying intent (curative vs. palliative) is not always made clear. Furthermore, many analyses or databases combine AS, WW, and noncurative treatments like primary ADT in their analyses, making it impossible to ferret out issues specifically related to AS.

The choice of immediate active treatment requires the careful consideration of a number of tradeoffs, such as balancing the harms of short- and long-term complications from curative treatments against the benefits of potential reductions in long-term morbidity and mortality. AS and other observational management strategies may, therefore, be considered by men who are more interested in avoiding the risks of curative treatment. Thus, it is important to clarify appropriate eligibility criteria and followup protocols for observational strategies that could minimize both unnecessary early curative treatments and avoidable prostate cancer symptoms and deaths. Of course, this strategy depends on the supposition that AS is as effective as (or no worse than) immediate curative treatments in an appropriate subgroup of men diagnosed with prostate cancer; this, however, remains to be proven. Furthermore, some men may be uncomfortable with observational management and feel a strong need to "do something," and

thus AS may be rarely offered, chosen, or adhered to. Therefore, the factors affecting these actions also warrant further investigation.

The National Cancer Institute (NCI) and the Centers for Disease Control and Prevention (CDC) are sponsoring a National Institutes of Health (NIH) State-of-the-Science Conference in December 2011 to examine the role of AS (as opposed to immediate curative intent therapy) in the management of early-stage, low-risk prostate cancer. The NIH has tasked the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) Program to provide an evidence review for use in this conference. The objective of this report is to summarize the existing literature on the role of AS in the management of early-stage, low-risk prostate cancer. Both the report and the corresponding NIH State-of-the-Science conference are a part of the NIH Consensus Development Program (CDP), the purpose of which is to evaluate the scientific evidence on a particular topic and develop a statement that advances research in this area. This statement is developed by an independent panel that is assembled for the conference. The panel will hear the scientific data, including the findings of this evidence review, and will then use that information to compose their statement. Additional information about the NIH CDP can be found at the NIH CDP Web site (http://consensus.nih.gov/).

The Conference planning committee crafted the Key Questions to be addressed at the conference, and the contracted EPC charged with systematically reviewing the literature to address them. Key Question 1 pertains to temporal trends in the natural history of prostate cancer in the United States. Key Question 2 relates to the definitions of observational (no active treatment) management strategies for prostate cancer used in the published literature. Key Question 3 relates to the factors that influence the offer or acceptance of or adherence to AS. Key Question 4 pertains to the comparative effectiveness of AS versus active treatments for localized prostate cancer. And Key Question 5 addresses recommendations for future research on observational management strategies for localized prostate cancer. It should be noted that this review primarily concerns active surveillance versus curative treatments. The widespread use of primary ADT in localized prostate cancer[5] is outside the scope of this review. The exact wording of the Key Questions to be addressed is provided below.

Key Questions

1. How have the patient population and the natural history of prostate cancer diagnosed in the United States changed in the last 30 years?
    a. Patient Characteristics
        i. Age
        ii. Comorbidity
        iii. Race/ethnicity
    b. Tumor Characteristics
        i. Stage
        ii. Tumor volume
        iii. Gleason score
        iv. PSA
    c. Diagnostic Strategies
        i. Biopsy Frequency
        ii. # of cores
        iii. Histopathologic grading changes
    d. System Characteristics

i. Differences in geographical access

2. How are active surveillance and other observational management strategies defined?
    a. Common Metrics
        i. Age
        ii. Gleason
        iii. # cores
        iv. % cores
        v. PSA (velocity, doubling time)
        vi. Imaging
        vii. Behavioral indicators
    b. Followup Protocols
        i. Gleason
        ii. # cores
        iii. % cores
        iv. PSA
        v. Imaging
        vi. Behavioral indicators
3. What factors affect the offer of, acceptance of, and adherence to active surveillance?
    a. Physician Factors
        i. Primary care
        ii. Diagnosing physician
        iii. Consultant – second opinion
        iv. Clinical factors
    b. Patient Factors
        i. Family involvement
        ii. Personal preferences
        iii. Risk perceptions
        iv. Family history
        v. Social support
    c. Delivery System
        i. Economic incentives and disincentives
            1. Insurance Type (HMO, military, private)
            2. Availability of technology
        ii. Geographic location
            1. Small area variation
            2. Regional variation
            3. Urban vs. rural
        iii. Academic centers vs. private practice
    d. Communication Strategies
        i. Risk assessment, predictive models
        ii. Decision-making tools and aids
4. What are the comparative short- and long-term outcomes of active surveillance versus immediate treatment with curative intent for localized prostate cancer?
    a. Prostate specific and all cause mortality
    b. Morbidity of primary treatment decision
    c. Incidence of metastatic disease

    d. Quality of life
    e. Costs
5. What are the research needs regarding active surveillance (or watchful waiting) in localized prostate cancer?

Methods

The present review evaluates trends in the epidemiology and natural history of prostate cancer in the United States. It also reports on aspects relevant to active surveillance (AS), watchful waiting (WW), and other "no treatment" approaches for managing localized disease. The evidence presented was obtained through a systematic review of the published scientific literature using established methodologies as outlined in the Agency for Healthcare Research and Quality's (AHRQ) Methods Guide for Comparative Effectiveness Reviews.[6]

AHRQ Task Order Officer

The Task Order Officer (TOO) was responsible for overseeing all aspects of this project. The TOO facilitated a common understanding among all parties involved in the project, resolved ambiguities, and fielded all EPC queries regarding the scope and processes of the project. The TOO and other staff at AHRQ reviewed the report for consistency, clarity, and to ensure that it conforms to AHRQ standards.

External Expert Input

The EPC convened a group of experts in the epidemiology and treatment of prostate cancer to form the Technical Expert Panel (TEP). Members of the TEP provided clinical and methodological expertise and input to help interpret the Key Questions guiding this review, identify important issues, and define parameters for the review of evidence. Discussions between the EPC, TOO, and the TEP occurred during a series of teleconferences and via email. In addition, input from the TEP was sought during compilation of the report when questions arose about the scope of the review. See Acknowledgement for the list of TEP members, and title page for our local domain experts.

Key Questions

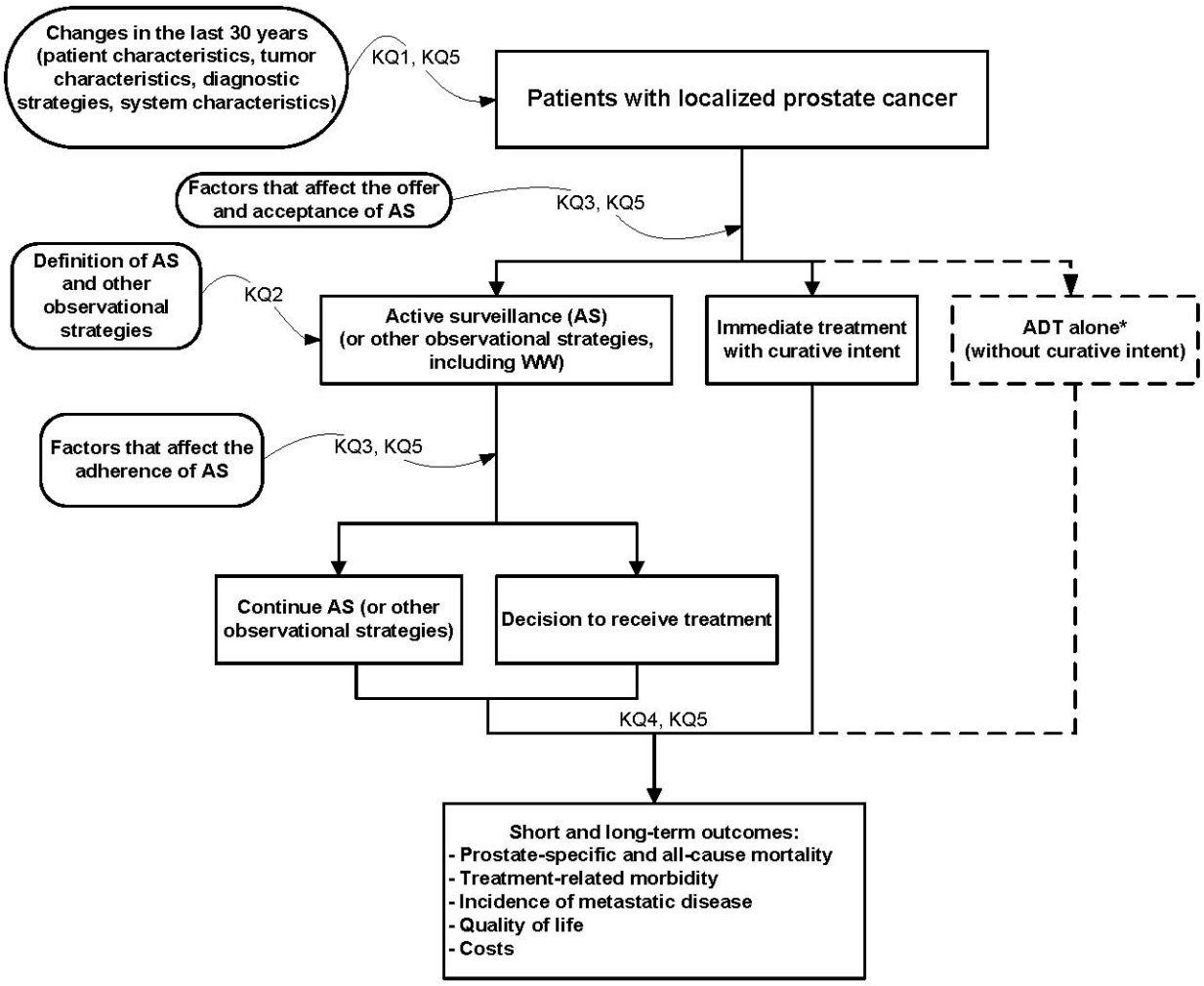
The Key Questions listed in the Introduction were provided by the NIH Consensus Development Program (CDP). The Key Questions have not been altered for the review.

Analytic Framework

To guide this systematic review and facilitate the interpretation of the Key Questions, we developed an analytic framework (Figure 1) that depicts the logical progression and interconnection of all five Key Questions of interest. The relevant population is patients with localized prostate cancer. Key Question 1 addresses changes in the last 30 years with respect to patient, tumor, and system-level characteristics at diagnosis, as well as trends in the diagnostic strategies employed. Key Question 2 examines the definitions of active surveillance (AS) and other observational strategies in terms of common metrics and followup protocols, as they have been implemented in clinical research. Patients diagnosed with localized prostate cancer are faced with a decision to either enter an AS monitoring protocol or receive immediate treatment with curative intent. Key Question 3 addresses the patient-, physician-, and system-level factors that influence this decision, in term of the offer and acceptance of, or adherence to AS. Key Question 4 addresses the short- and long-term outcomes and costs associated with AS versus immediate treatment with curative intent. Outcomes of interest include prostate specific- and all-cause mortality, morbidity of primary treatment, incidence of metastatic disease, quality of life,

and cost. Key Question 5 addresses future research needs across the spectrum of Key Questions 1 through 4.

Figure 1. Analytic framework that depicts the five Key Questions (KQs) that examine the role of active surveillance in the management of men with clinically localized prostate cancer



ADT = androgen deprivation therapy; AS = active surveillance, KQ = key question.
* ADT alone is a management strategy that is not reviewed in this review.

Literature Searches

Studies included in this review were identified through multiple literature searches using terms relevant to prostate cancer or observational management strategies (including AS and WW) (provided in Appendix A). Specifically, we performed a search for systematic reviews and meta-analyses in the MEDLINE database (from 1996 through December week 4, 2010). We did not search for systematic reviews and meta-analyses published earlier to ensure the results would be applicable to current clinical practice. We performed an additional search of the MEDLINE database (from inception through August, 2011) using terms for specific databases (such as the Surveillance Epidemiology and End Results (SEER) database and the Cancer of the Prostate

Strategic Urologic Research Endeavor (CaPSURE) database) along with terms for prostate cancer. This search was supplemented by a MEDLINE search (from inception through August, 2011) combining terms relevant to observational management strategies (e.g., WW, AS, expectant management) along with terms for prostate cancer; the search strategy was based on expanding a previously published set of keywords.[7] We also performed a targeted literature search for economic evaluations using the Cost-Effectiveness Analysis Registry[a], a curated database of cost-effectiveness analyses maintained by the Center for the Evaluation of Value and Risk in Health at Tufts Medical Center (the search was performed on November 2, 2011; at that time the registry contained data up to 2010).

Additional citations were provided by members of the TEP and were reviewed against the same inclusion and exclusion criteria that were used for studies identified through the database searches. We also perused the reference lists of the eligible primary studies and relevant review articles to identify additional potentially relevant studies.

For the question (Key Question 3) related to factors affecting AS, we did not do a targeted search of the individual factors, instead we relied on primary studies and systematic reviews already identified via the above searches.

We did not include unpublished data (such as abstracts or meeting proceedings) for this review.

After running each search, abstracts were entered in an electronic database and nonoverlapping sets of citations were screened by a single investigator.

We identified ongoing studies (provided in Appendix B) through perusal of reference lists, review articles and citations from the TEP.

Study Selection and Eligibility Criteria

Based on input from the TEP and the TOO, we developed selection criteria for identifying studies for each Key Question. These criteria were different for systematic reviews and primary research studies and are summarized below. For all Key Questions we excluded editorials, letters to the editor, narrative reviews, and any other publications not presenting research results or describing the protocols of primary research studies. We only considered English language studies.

Systematic Reviews, Evidence Reports, Economic Evaluations

We defined systematic reviews as studies using explicit methods to search, identify and synthesize primary research studies. Reviews utilizing both qualitative or quantitative (meta-analysis) methods to synthesize the available evidence were considered eligible, as long as they provided information considered relevant to the Key Questions.

For Key Question 1, in addition to systematic reviews we also reviewed the latest version of the Cancer Statistics report prepared annually by the American Cancer Society[1], a recent SEER Survival Monograph[2], the 2009/2010 update of the Cancer Trends Progress Report[3], and data available on the SEER Web site.[b]

For Key Question 2, we only used the reference lists of relevant systematic reviews to identify additional primary studies providing definitions of observational management strategies.

[a] Available at https://research.tufts-nemc.org/cear4/SearchingtheCEARegistry/SearchtheCEARegistry.aspx; last accessed October 31, 2011.

[b] Available at http://seer.cancer.gov/faststats/; last accessed September 30, 2011.

For Key Question 3, we reported relevant findings from existing systematic reviews and we also perused the reviews' references to identify potentially eligible studies.

For Key Question 4 we mainly relied on two AHRQ evidence reports.[8,9] For treatment-related costs, we considered economic evaluations (cost components of cost-effectiveness or cost-utility analyses, and model-based cost-minimization or cost-consequence analyses) of AS only. We did not include the effectiveness components of cost-effectiveness or cost-utility analyses because they did not provide additional comparative treatment information beyond that provided by the AHRQ reports and our own literature searches.

For all Key Questions, when necessary, the evidence summarized in previously published systematic reviews and evidence reports was supplemented with studies identified through our own literature searches

Primary Research Studies

Key Question 1 (Trends in Incidence, Mortality/Survival and Features at Diagnosis)

For Key Question 1 we included studies utilizing large databases sourced from the U.S. population (e.g., the SEER database or its component registries; the CaPSURE database; or the National Cancer Database [NCDB]). We excluded studies conducted in other countries. We required the reported patient data to be within the time period 1980 to 2011. Studies had to have analyzed data from at least 1000 patients and to report numerical data informing on changes of the parameters of interest (incidence, mortality/survival, patient-, tumor- or system-level characteristics at diagnosis, treatment patterns).

We required that studies reported changes over time or stratified by time periods (with or without an associated statistical test). Thus, we excluded studies of single years. We included studies that treated time as continuous variable (e.g., year of diagnosis) or as a categorical variable (e.g., "before 2000" versus "after 2000"). Studies reporting only qualitative descriptions were excluded. We also excluded studies that only reported on prostate cancer patients who were selected based on the treatment modality they received (e.g., we did not include studies where patients had to have received radical prostatectomy or studies excluding patients receiving AS/WW), studies enrolling patients exclusively diagnosed through transurethral resection of the prostate performed for benign prostatic hyperplasia, and single-center studies. The latter were excluded to maximize the applicability of the included studies to the U.S. population.

Because differences in patient selection criteria or underlying populations can confound temporal trends, we only considered studies where trend data were obtained from within the same database, and we avoided inferences on temporal trends across manuscripts or databases.

Key Question 2 (Definitions of Observational Management Strategies)

We considered studies reporting on observational management strategies (i.e., no immediate active treatment with curative intent), enrolling patients based on predefined eligibility criteria, and using prespecified protocols for followup. We considered both studies where the aim of observational management was to offer curative treatments when disease progression meets predefined laboratory and clinical parameters in a monitoring protocol (AS) and studies where the aim of observational management was to offer palliative treatments when patients become clinically symptomatic (WW). Both prospective and retrospective studies of any design were

considered eligible. We evaluated only the descriptions of the observational management strategies.

When a center or research group had published multiple studies reporting on potentially overlapping patient populations, the publication that provided the most complete information on eligibility criteria and followup protocols (i.e., the study that was most informative regarding the components relevant to Key Question 2) was used as the primary source of information for this report. We also considered additional publications from the same cohort when they reported important changes (e.g., in cases where papers explicitly reported changes in the study protocol that affected the definition of the observational strategy). When all articles from the same center or research team used the same observational strategy (i.e., when the same definition was consistently used in all publications), we generally referenced the article with the earliest publication date. We included studies from any country.

Key Question 3 (Factors Affecting the Offer of, Acceptance of, and Adherence to Observational Management Strategies)

We considered three types of studies relevant to Key Question 3: (1) studies of any design that used multivariable methods to predict the offer, acceptance, or adherence of observational management; (2) studies that used qualitative research methods to identify such factors; and (3) experimental studies that examined a factor of interest addressing the same issues, when applicable (e.g., decision aids). We included studies from any country. For each type of study we employed different criteria:

Studies Using Quantitative Methods to Predict Offer, Acceptance or Adherence

For this category, we considered studies reporting on factors predicting the offer of, acceptance of, or adherence to observational management strategies (including AS and WW). We excluded studies explicitly reporting that patients receiving androgen deprivation therapy (ADT) were considered together with patients receiving no treatment (WW or AS). We also excluded studies defining observational management strategies as the lack of surgical treatment or radiotherapy, without providing information on how other treatments (such as ADT) were handled in the analyses. However, studies that provided no definition of the observational strategy employed (e.g., "expectant management" with no other information on how treatment groups were defined) were included.

Both prospective and retrospective studies of any design were eligible, so long as one treatment group was managed using observational management strategies. We required that studies used multivariable methods (e.g., multivariable regression or analysis of covariance) to adjust for potential confounders. At a minimum we required adjustment for age and tumor stage (if the analysis was not limited to patients with localized cancer).

Studies Using Qualitative Methods

For this category, we considered studies using qualitative research methods (e.g., focus groups or surveys) to obtain information on factors that affect the offer of, acceptance of, or adherence to AS or WW. Eligible studies had to use a predefined approach to collect information (e.g., a structured or semi-structured interview, a questionnaire).

Experimental Studies

For this category, we included studies of any design that evaluated any tool (such as a decision aid) or other intervention designed specifically to have an impact on acceptance of AS.

Key Question 4 (Comparative Effectiveness of Observational Management Strategies and Active Treatment)

We considered studies that fulfilled the following criteria:

Population. Men with clinically localized prostate cancer (T1-T2), without (or unable to assess) either regional lymph nodes involvement (N0-X) or metastases (M0-X), regardless of age, histologic grade, Gleason score, or prostate-specific antigen (PSA) concentration. Studies that enrolled mixed populations of clinically localized and more advanced disease were included only if men with more advanced disease stages represented less than 20 percent of the study sample, or if they reported separate treatment effect estimates for the clinically localized subgroup of patients.

Intervention. Observational (no immediate active treatment) management strategies, including both WW and AS strategies. We excluded studies where the observational management group was combined with the group of patients receiving either medical or surgical ADT (i.e., orchiectomy). We also excluded studies defining observational management strategies as the lack of surgical treatment or radiotherapy, without providing information on how other treatments (such as ADT) were handled in the analyses. However, studies that provided no definition of the observational strategy employed (e.g., "expectant management" with no other information on how treatment groups were defined) were included.

Comparators. Radical prostatectomy (RP), external beam radiation therapy (EBRT), or brachytherapy (BT), all with or without ADT. Based on input from the TEP and AHRQ representatives, ADT was not considered a potentially curative treatment and studies that compared observational management strategies only with ADT monotherapy were excluded.

Outcomes. Outcomes of interest included prostate cancer-specific mortality, all-cause mortality, morbidity of primary treatment (including the frequency of procedures to address treatment-related morbidity), development of metastatic disease, quality of life (QoL, including satisfaction with treatment), and costs. Eligible studies had to report or provide sufficient data to allow the estimation of the treatment effect (e.g., hazard ratios, odds ratios, risk differences, or risk ratios along with sufficient statistics to calculate the uncertainty around these estimates) or provide the P value from a test of association of the treatments examined with the outcomes of interest. For treatment-related costs, we considered estimates from primary studies of actual patient costs (i.e., based on data collected during patient followup) of any observational management strategy (AS and WW).

Study design and analysis. We considered both randomized controlled trials, and prospective or retrospective nonrandomized comparative studies with longitudinal followup from any country. We excluded cross-sectional and case-control studies. Eligible observational studies had to be conducted in a multicenter setting in any country or to have utilized databases sourced from the U.S. population (such as SEER, CaPSURE, the Prostate Cancer Outcomes Study [PCOS], the

Patterns of Care Study [POCS], or Veterans Administration studies). Nonrandomized comparative studies also had to use multivariable methods (regression or propensity-score based) or instrumental variable methods to estimate treatment effects. Operationally, we required adjustment at least for patient age for all observational studies; when such studies reported on mixed populations (localized mixed with more advanced disease), we also required adjustment for at least one marker of disease severity (e.g., disease stage, tumor grade, Gleason score).

For all Key Questions, potentially eligible studies identified through screening titles and abstracts (see previous section) were retrieved in full text and were reviewed by a single investigator using the above listed criteria.

Key Question 5 (Research Needs)

We did not perform a separate literature search for this Key Question, but instead reviewed the evidence for Key Questions 1 to 4 to identify research gaps.

Data Extraction and Summaries

We considered primary research studies of diverse designs, including published systematic reviews of primary research studies. We list here the information that was extracted from these types of evidence. The summary Tables are provided in Appendix C.

Primary Research Studies

We extracted bibliographic information, eligibility criteria, enrollment years, study duration, and the number of patients included in the final analytic sample. We also extracted additional information from primary research studies considered relevant to each Key Question.

Key Question 1 (Trends in Incidence, Mortality and Features at Diagnosis)

From each study that provided information on temporal trends we extracted information that allowed the reconstruction of trends over time in incidence, mortality/survival, patient-, tumor-, and system-level characteristics at diagnosis. For parsimony, we grouped the extracted information in 5-year bins covering the time period of interest (1980-2010/11). When a study reported multiple estimates of the parameters of interest within a single 5-year bin, we only extracted information for the year closest to the mid-point of the bin (e.g., if a study reported incidence rate data for all years between 1980 and 1985, we extracted the incidence rates for the years 1982 only).

From studies reporting statistical tests for change of the parameters of interest over time, we extracted the following information (when available): the specific method used for statistical analysis of trend data, estimates of trend statistics, and p-values for changes in parameters of interest over time.

Key Question 2 (Definitions of Observational Management Strategies)

To describe the definitions of observational management strategies used in published studies of such strategies, we extracted information on patient- and tumor-level characteristics used as eligibility criteria, followup or monitoring parameters, or specific triggers for intervention (active therapy). We also extracted details on the definition of disease progression used in each study.

We took particular care to identify changes in the observational protocols used by research teams that had published more than one paper providing information relevant to Key Question 2.

Key Question 3 (Factors Affecting the Offer, Acceptance and Adherence to Observational Management Strategies)

For studies using multivariable models to identify factors associated with the offer, acceptance or adherence of WW or AS, we extracted information on the definition of the observational strategy evaluated in each study, the statistical analysis methods used to identify factors of interest, and the main findings as related to Key Question 3.

For studies using qualitative methods to identify factors associated with the offer, acceptance or adherence of WW or AS, we extracted information on the research methods used, the definition of the observational strategy addressed in each study, and a qualitative summary of key study findings.

Key Question 4 (comparative effectiveness of observational management strategies and active treatment)

From each eligible comparative treatment study we extracted the following information: detailed descriptions of the interventions being compared, the source populations of each study, details of the eligibility criteria used, sample size information, study start and end dates, followup duration, baseline characteristics of the enrolled patient populations, measurement instruments, the definitions of specific outcomes, and estimates of the treatment effect.

Systematic Reviews

For systematic reviews, we extracted information on the data sources used, the dates covered by the literature searches, the inclusion and exclusion criteria used, the number of eligible studies identified, whether quantitative synthesis (meta-analysis) was performed, and a description of key study findings.

Quality Assessment

Primary Research Studies

We assessed the methodological quality of only observational and randomized studies included for Key Question 4. The EPC, in consultation with the TOO, decided that formal quality assessment was unlikely to be informative for Key Questions 1-3 because it was not deemed well applicable to the descriptive literature summarized for these Key Questions. For Key Question 4, quality assessment was performed by the team member doing the primary data extraction. The quality grade was confirmed by at least one other team member.

We assessed the methodological quality of studies based on predefined criteria. We used a three-category grading system (A, B, or C) to denote the methodological quality of each study as described in the AHRQ methods guide.[6] This grading system has been used in most of the previous evidence reports generated by our EPC. This system defines a generic grading scheme that is applicable to varying study designs including RCTs, nonrandomized comparative trials, cohort, and case-control studies. For RCTs, we primarily considered the methods used for randomization, allocation concealment, and blinding as well as the use of intention-to-treat analysis, the report of dropout rate, and the extent to which valid primary outcomes were

described as well as clearly reported. Only RCTs and prospective comparative studies could receive an A grade. Retrospective studies could be graded either B or C. For all studies, we used (as applicable): the report of eligibility criteria, the similarity of the comparative groups in terms of baseline characteristics and prognostic factors, the report of intention-to-treat analysis, important differential loss to followup between the comparative groups or overall high loss to followup, and the validity and adequacy of the description of outcomes and results.

A (good). Quality A studies have the least likelihood of bias, and their results are considered most valid. They generally possess the following: a clear description of the population, setting, interventions, and comparison groups; appropriate measurement of outcomes; appropriate statistical and analytic methods and reporting; no reporting errors; clear reporting of dropouts and a dropout rate less than 20 percent; and no obvious bias. Only prospective studies may receive a grade of A.

B (fair/moderate). Quality B studies are susceptible to some bias, but not sufficiently to invalidate results. They do not meet all the criteria in category A due to some deficiencies, but none likely to introduce major bias. Quality B studies may be missing information, making it difficult to assess limitations and potential problems.

C (poor). Quality C studies have been adjudged to carry a substantial risk of bias that may invalidate the reported findings. These studies have serious errors in design, analysis, or reporting and contain discrepancies in reporting or have large amounts of missing information.

Systematic Reviews

Our assessment of systematic reviews was based on methodological guidelines for reviews of studies of therapeutic interventions[10] or epidemiological studies.[11] We also assessed the quality of reviews by extracting information on the items included in the Assessment of Multiple Systematic Reviews (AMSTAR) checklist.[12,13] Because AMSTAR was developed for typical published systematic reviews (and not evidence reports where often the use of specific methods or reporting practices is not at the discretion of the investigators) we did not use this checklist to assess the quality of evidence reports considered as sources of evidence for this review (i.e., the AHRQ reviews on prostate cancer treatments).

Data Synthesis and Presentation

We summarized all included studies in narrative form as well as in summary tables (see below) that condense the important features of the study populations, design, intervention, outcomes, and results. For Key Questions 1-4 we synthesized the extracted information qualitatively. Because there was extensive heterogeneity in reporting information for most variables and substantial potential for population overlap between studies for all Key Questions (e.g., the majority of epidemiologic studies considered eligible for Key Question 1 were based on the SEER and CaPSURE databases and covered overlapping periods of time), we did not perform additional quantitative analyses (meta-analyses).

When appropriate we summarized the characteristics of eligible studies using summary statistics (means, medians, ranges and standard deviations).[14] For Key Question 1, we created line graphs depicting trends over time using publicly available information from the SEER Web site (http://seer.cancer.gov/; last accessed September 30, 2011). For Key Question 2, we

generated bar graphs showing the number of AS cohorts employing each specific criterion for patient selection or as part of their followup protocol, to demonstrate items for which heterogeneity was most prominent across cohorts.

Summary Tables

Summary tables succinctly report measures of the main outcomes evaluated. We included information regarding population selection, country (when relevant), study design, interventions, age data, study setting, prostate cancer stage and grade, sample size, study duration, years of intervention, dropout rate, and study quality (for Key Question 4). For continuous outcomes, we included the mean outcome values, their 95 percent confidence intervals (CI), standard deviations (SD) or other measures of variability and when available, the mean difference (between groups) and its corresponding P value, or CI, as appropriate. For categorical (dichotomous) outcomes, we reported the number of events and total number of patients for each intervention and relative risk metrics (odds ratios, risk ratios or hazard ratios) with their corresponding 95 percent CI and associated P value.

Grading the Body of Evidence for Key Question 4

We graded the strength of the body of evidence for each analysis within Key Question 4 as per the AHRQ methods guide[6] and an updated methods paper,[15] with modifications as described below. Risk of bias was assessed using a three-category grading system (A, B, or C) which corresponds to low, medium or high risk of bias (see Quality Assessment). We assessed the consistency of the data as either "no inconsistency" or "inconsistency present" (or "not applicable" if only one study). The direction, magnitude, and statistical significance of all studies were evaluated in assessing consistency, and logical explanations were provided in the presence of equivocal results. Studies with limited relevance either included populations which related poorly to the general population of men in the United States with localized prostate cancer or contained substantial problems with the measurement of the outcome(s) of interest. We also assessed the precision and sparseness of the evidence. We considered evidence to be sparse if only one study addressed the analysis.

We rated the strength of evidence with one of the following four strengths (as per the AHRQ methods guide): High, Moderate, Low, and Insufficient. Ratings were assigned based on our level of confidence that the evidence reflected the true effect for the major comparisons of interest. Ratings were defined as follows:

High. There is high confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect. No important scientific disagreement exists across studies. At least two quality A studies are required for this rating. In addition, there must be evidence regarding objective clinical outcomes.

Moderate. There is moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate. Little disagreement exists across studies. Moderately rated bodies of evidence contain fewer than two quality A studies or such studies are inconsistent or lack long-term outcomes of relevant populations.

Low. There is low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate of effect and is likely to change the estimate. Underlying studies may report conflicting results. Low rated bodies of evidence could contain either quality B or C studies.

Insufficient. Evidence is either unavailable or does not permit a conclusion. There are sparse or no data. In general, when only one study has been published, the evidence was considered insufficient, unless the study was particularly large, robust, and of good quality.

These ratings provide a shorthand description of the strength of evidence supporting the major questions we addressed. However, they by necessity may oversimplify the many complex issues involved in appraising a body of evidence. It is important to remember that the individual studies involved in formulating the composite rating differed in their design, reporting, and quality. The strengths and weaknesses of the individual reports, as described in detail in the text and tables, should also be taken into consideration.

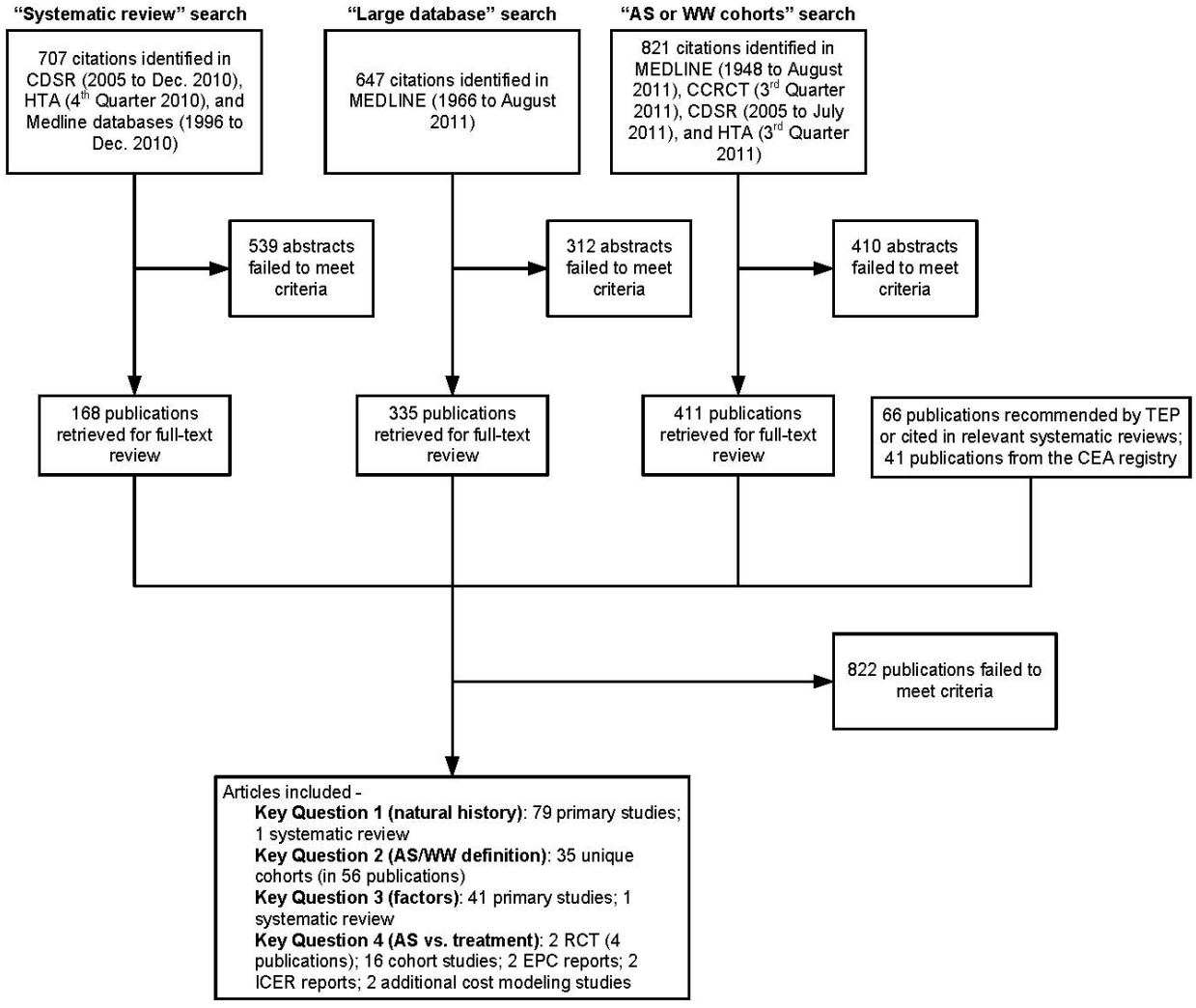
Peer Review and Public Commentary

As part of a newly instituted process at AHRQ, the initial draft report was prereviewed by the TOO and an AHRQ Associate Editor (a senior member of a sister EPC). Following revisions, the draft report was sent to invited peer reviewers and was simultaneously uploaded to the AHRQ Web site where it was available for public comment for 30 days. All reviewer comments (both invited and from the public) were collated and individually addressed. The authors of the report had final discretion as to how the report was revised based on the reviewer comments, with oversight by the TOO and Associate Editor.

Results

Our literature searches for prostate cancer systematic reviews, for large registry databases, and for AS/WW publications yielded 707, 647, and 821 citations, respectively. From these, 914 articles were provisionally accepted for review based on the abstracts and titles. Additional citations recommended by technical expert panel, from reference lists of relevant systematic reviews, or identified through targeted searches of the Cost-Effectiveness Analysis Registry, were also accepted for review. After full-text screening, 195 papers and 4 evidence/economic reports met criteria and are included in the review (Figure 2).

Figure 2. Literature flow



The numbers of studies for each Key Question do not sum to the total number of included studies because some studies addressed multiple Key Questions. AS = active surveillance; CDSR = Cochrane Database of Systematic Reviews; HTA = Health Technology Assessment; CCRCT = Cochrane Central Register of Controlled Trials; CEA Registry = cost effectiveness analysis registry; EPC = Evidence-based Practice Center; ICER = Institute for Clinical and Economic Reviews; TEP = technical expert panel; RCTs = randomized controlled trials; WW = watchful waiting.

Key Question 1. How have the patient population and the natural history of prostate cancer diagnosed in the United States changed in the last 30 years?

Prostate cancer epidemiology is affected by population-level trends, such as the aging of the United States population, but also by changes in the application of screening and diagnostic technologies among the population at risk. To assess temporal trends in the incidence, mortality/survival, disease features at diagnosis, and treatment patterns we performed a search to identify large studies (≥1000 men) utilizing databases sourced from the U.S. population that provided information stratified by factors relevant to Key Question 1 (see the end of the Introduction for the list of factors). We also reviewed the latest version of the Cancer Statistics report prepared annually by the American Cancer Society[1], a recent SEER Survival Monograph[2], the 2009/2010 update of the Cancer Trends Progress Report[3], and data available on the SEER Web site.[a]

We identified 79 primary observational studies and one systematic review eligible for inclusion in Key Question 1.[4,5,16-92]

Of the primary observational studies, 51 analyzed the Surveillance Epidemiology and End Results (SEER) database of the National Cancer Institute (NCI) or a subset of its component registries, nine additional studies utilized the linked SEER-Medicare database, 11 the Cancer of the Prostate Urologic Research Endeavor (CaPSURE) database, five the National Cancer Database (NCDB), and three analyzed other large U.S.-based databases.

- The SEER database consists of a coordinated system of population-based cancer registries covering geographic areas selected for inclusion based on their ability to provide high quality population-based cancer reporting and for their epidemiologically significant population subgroups. The SEER population is comparable to the general U.S. population with regard to measures of poverty and education; however, the SEER population tends to be somewhat more urban and has a higher proportion of foreign-born persons than the general U.S. population.[b]
- The racial and age distribution on SEER areas is also not perfectly representative of the total U.S. population and the data may be insufficient for minority groups other than blacks.[93,94]
- The SEER-Medicare database linked the SEER cancer registries data and Medicare enrollment and claims files.[c]
- A comparison of sociodemographic characteristics of Medicare beneficiaries residing in the SEER areas versus the general U.S. elderly population demonstrated that the age and sex distribution for individuals 65 years and older in the SEER areas was comparable to that of the U.S. elderly population. However, the elderly population in the SEER areas had a lower proportion of whites and a higher proportion of other racial/ethnic groups and was also more likely to reside in an urban setting compared with the average 65 years and older U.S. population.[95]
- The CaPSURE database includes data from a longitudinal, observational study of over 14,000 men with all stages of biopsy-proven prostate cancer. Patients are enrolled regardless of age, stage of disease, or intended treatment plan. Currently, CaPSURE

[a] Available at http://seer.cancer.gov/faststats/; last accessed September 30, 2011.
[b] http://seer.cancer.gov/registries/characteristics.html ; last accessed September 30, 2011.
[c] See http://healthservices.cancer.gov/seermedicare/; last accessed September 30, 2011.

collects data from 40 urology practices in the United States (34 community based, 3 Veterans Administration [VA])-based and 3 academic center based).[d] Although CaPSURE patients are broadly representative of prostate cancer diagnosed in the community, patient selection is not based on a formal population sampling scheme.

- The NCDB is an oncology outcomes database encompassing more than 1,500 cancer programs in the U.S. (including Puerto Rico) accredited by the Commission on Cancer of the American College of Surgeons and the American Cancer Society.[e]
- Other databases utilized by the primary studies were the Patterns of Care study (sponsored by NCI and based on sampling participants through SEER) and the Los Angeles County/University of Southern California (LAC/USC) Cancer Surveillance Program (which is now a component registry of SEER).

In addition to the above databases, several studies obtained prostate cancer mortality data from the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC), the U.S. principal health agency for vital statistics.[f]

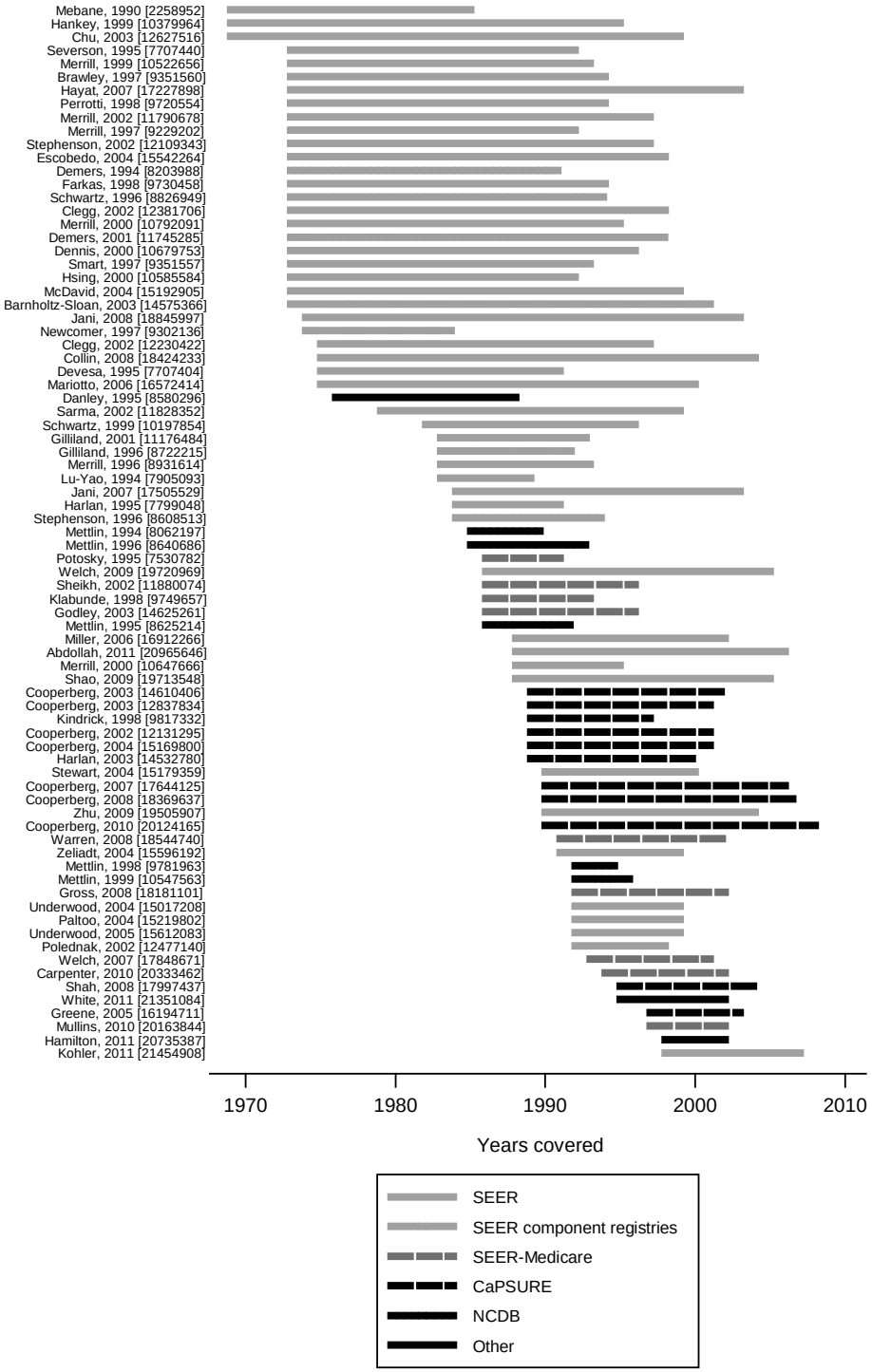
Included studies had large sample sizes (median sample size = 54,670 patients; 25^{th}-75^{th} percentile 11,161-138,387), were published between 1990 and 2011, and analyzed data from 1969 to 2008. Figure 3 presents the years covered by each primary study and the databases used. Appendix Tables C1.1-C1.14 present additional information about each of the studies relevant to Key Question 1.

We organized the 79 studies into four groups, each of which is discussed in the following sections:

1. studies investigating trends in prostate cancer incidence
2. studies investigating trends in prostate cancer mortality or survival
3. studies investigating patient-, tumor-, or system-level characteristics at prostate cancer diagnosis, and
4. studies presenting information in treatment trends over time

[d] See http://urology.ucsf.edu/clinicalres/CRuroOnc_gceps_capsure.html; last accessed September 30, 2011.
[e] See http://www.facs.org/cancer/ncdb/index.html; last accessed September 30, 2011.
[f] See http://www.cdc.gov/nchs/; last accessed September 30, 2011.

Figure 3. Years covered and databases utilized by studies considered eligible for Key Question 1

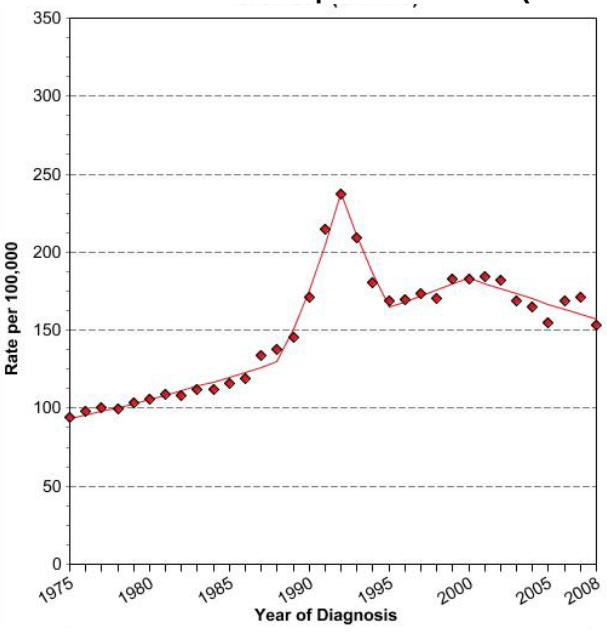


Horizontal lines indicate the years covered by each primary study considered for Key Question 1. Different line patterns indicate the different databases utilized by each study. Studies are listed by the first year covered, then by database used, then by year of publication and are presented using the format: first author, year of publication [MEDLINE unique identifier]. Though data from earlier years were available, we analyzed only data from 1980 onward. Studies using SEER along with other information sources have been grouped in the "SEER" category for simplicity.

Trends in Prostate Cancer Incidence

Prostate cancer incidence trends in the United States during the last 30 years have been largely driven by changes in screening practices, mainly the implementation of prostate specific antigen (PSA) screening. Empirical analyses of incidence data and simulation studies demonstrate that patterns in prostate cancer incidence are compatible with the introduction and widespread use of a sensitive screening test, resulting in increases in the number of new cases diagnosed every year. The NCI's Cancer Trends Progress Report (2009/10) indicates that prostate cancer incidence rates rose between 1975 and 1992 (from approximately 100 to more than 240 new cases per 100,000 men per year), and then fell until around 1995. After a period of nonsignificant increase from 1995 to 2000, rates declined again from 2000 to 2007 (to the current level of approximately 156 new cases per 100,000 men per year; Figure 4).

Figure 4. Age-adjusted SEER incidence rates for prostate cancer (1975–2008)



Only includes invasive cancer cases. Incidence data pertain to SEER9 areas: San Francisco, Connecticut, Detroit, Hawaii, Iowa, New Mexico, Seattle, Utah and Atlanta. Rates are presented per 100,000 and are age-adjusted to the 2000 U.S. standard population. Regression lines were fitted through joinpoint regression. Image obtained from SEER Fast Stats (http://seer.cancer.gov/faststats/index.php; last accessed September 30, 2011).

Overall, 33 studies provided information on trends in prostate cancer incidence (Appendix Table C1.2). Of these, 11 provided information stratified by patient age,[17,27,28,35,38,39,43,55,57,59,79] 17 by race/ethnicity,[16,17,19,21,25,32,34,42,43,45,46,49,57,76,80,83,84] 15 by tumor stage,[17,18,20,23,26,27,34,35,42,50,52,55,56,76,85] and 5 by tumor grade.[17,21,42,52,56] No studies that met our inclusion criteria provided information stratified by the other factors relevant to Key Question 1 (comorbidity, tumor volume, PSA, biopsy frequency, number of cores obtained at biopsy, or system level characteristics).

Studies providing information on cancer incidence were large (median sample size = 51,337; 25th-75th percentile 39,566-100,212), were published between 1990 and 2011, and provided information for years 1969 to 2007.

Patient Characteristics

Age. Eleven studies (9 SEER, 2 SEER-Medicare) covering 1969 to 2005 provided information on prostate cancer incidence stratified by patient age.[17,27,28,35,38,39,43,55,57,59,79] Generally, studies indicated that since the mid-1980s the prostate cancer incidence rate increased across all age groups until 1992-93 and then declined until 1995-99. Data for more recent years were sparse. However, based on a study utilizing the SEER database, compared to the pre-PSA era (1986), the incidence rates in 2005 were 3.64 times higher for men aged 50-59 years (95 percent CI 6.4-8.2), 1.91 times higher for men aged 60-69 (95 percent CI 1.8-2.0), 1.09 times higher for men aged 70-79 years (95 percent CI 1.05-1.14), but 0.56 times less common for men 80 years or older (95 percent CI 0.53-0.60).[38]

Race/ethnicity. Seventeen studies (15 SEER, 1 SEER-Medicare, 1 LAC/USC) covering 1973 to 2007 provided information on prostate cancer incidence stratified by patient race/ethnicity.[16,17,19,21,25,32,34,42,43,45,46,49,57,76,80,83,84] Thirteen of the 17 studies provided information exclusively for whites or blacks and only four provided information on patients belonging to other racial or ethnic groups; data for other racial/ethnic groups were often only provided in aggregate (not separately for each ethnic group). Studies indicated that all racial/ethnic groups experienced increases in prostate cancer incidence since the mid-1980s. The incidence rate appears to have peaked in 1992 for non-Hispanic whites, in 1993 for blacks, and in 1992 for "other" racial/ethnic groups. For all groups, incidence rates declined between the early-1990s and 1999. One study that provided information up to 2005, reported that the incidence rates in recent years are higher compared to the pre-PSA era but lower than the peak values reached in the mid-1990s both for whites and blacks (both races $P < 0.001$ for the increase from 1988-89 to 2004-05).[42]

Tumor Characteristics

Stage. Fifteen studies (14 SEER, 1 LAC/USC) covering 1969 to 2005 provided information on prostate cancer incidence stratified by tumor stage at diagnosis.[17,18,20,23,26,27,34,35,42,50,52,55,56,76,85]

Fourteen studies investigated trends in the incidence of localized/regional and distant disease. Generally, studies consistently demonstrated that early-stage (localized and regional; several studies did not report data separately for these stages) prostate cancer cases were responsible for the observed increase in prostate cancer incidence from the mid-1980s up to the mid-1990s. When separate estimates were available, studies reported large increases in the incidence of localized disease after the introduction of PSA testing. Over the same period, studies demonstrated decreases in distant (metastatic) prostate cancer incidence. For example, following the introduction of PSA screening, a study using the SEER-Seattle-Puget Sound registry demonstrated a 60 percent decrease in the age-adjusted incidence rate of distant prostate cancer ($P < 0.001$ comparing 1986 to 1991).[55] Studies also consistently demonstrated decreases in incidence rates for all disease stages from mid-1990s to 2000. No study reported relevant information after 2000.

A single study (analyzing SEER) investigated changes in the distribution of T stage over time and demonstrated that compared to 1988-89, in 2004-05 the incidence rate had increased by 76 cases per 100,000 person-years for T1 tumors and by 11.2 cases per 100,000 person-years for T2 tumors. In contrast, over the same time period, the incidence of T3 or T4 tumors (combined) had decreased by 47.1 cases per 100,000 person-years ($P < 0.001$ for the stage specific changes).[42]

Tumor grade. Five studies (all using the SEER database or its component registries) covering 1973 to 2005 stratified prostate cancer incidence by tumor grade (level of differentiation or Gleason score).[17,21,42,52,56] Studies generally indicated that the increase in prostate cancer incidence observed from the mid-1980s to early-1990s was mainly due to an increase in the incidence rate of moderately differentiated tumors (or tumors of Gleason score 5-7). A single study (of SEER) analyzed data after 2000 and reported a continued increase in incidence rate of tumors with Gleason score 5-7 from 1988 to 2005 and a concomitant decrease in the incidence rate of tumors with Gleason score 2-4 ($P < 0.001$ for Gleason group-specific changes).[42]

Trends in Prostate Cancer Mortality and Survival Rates

For the overall U.S. population, the NCI's Cancer Trends Progress Report (2009/10) indicates that after increasing from 1975 to 1991, prostate cancer death rates fell from 1994 to 2007.[a] The baseline prostate cancer mortality in 1975 was 31 deaths per 100,000 men per year and has declined to the current level of approximately 24 deaths per 100,000 men per year.

Among the studies we reviewed, 21 provided information on trends in prostate cancer mortality or changes in survival rates of patients with prostate cancer, stratified by the factors relevant to Key Question 1 (Appendix Table C1.3). Nine of the studies provided information stratified by age,[17,18,24,28,30,33,39,40,86] 15 by race/ethnicity,[16-19,24,26,27,34,40,44,51,60,76,84,86] one by comorbidity status,[24] six by tumor stage,[18,24,40,78,81,86] and three by tumor grade.[24,40,86] No studies provided information stratified by any of the other factors relevant to Key Question 1. Of the studies considered eligible, 19 utilized the SEER database, one the SEER-Medicare database and one LAC/USC. Studies were generally large (median sample = 60,494; 25^{th}-75^{th} percentile 40,548-229,556), were published between 1990 and 2011, and covered years 1969 to 2007.

Patient Characteristics

Age. Nine studies (all SEER-based) covering 1969 to 2004, reported information on population mortality rates (5 studies) or prostate cancer-specific survival among patients diagnosed with prostate cancer survival (4 studies), stratified by age.[17,18,24,28,30,33,39,40,86] Studies of population mortality rates demonstrated decreases in the mortality rate for all age groups between the early-1990s and 1999. One study presented information for years after 2000 and showed that population mortality continued to decline significantly for all age groups in recent years (up to 2004).

One study of prostate cancer-specific survival, using the SEER database, demonstrated that over time (for deaths occurring from 1988 to 1995) the proportion of patients diagnosed with prostate cancer who died of their cancer has decreased (i.e., patients with prostate cancer have increasingly died of other causes) across all age groups considered (> 50 years old).[40] Another study, also using the SEER database, demonstrated that among patients with prostate cancer, increasing age is associated with death from nonprostate cancer causes and that this effect held true throughout the study period.[24]

Comorbidity. A single study using the SEER database reported information on temporal trends in prostate cancer mortality stratified by whether patients had been diagnosed with multiple primary cancers.[24] The study demonstrated that throughout the study period (1988 to 1995)

[a] Available at http://progressreport.cancer.gov/; last accessed September 30, 2011.

prostate cancer patients with multiple primary cancers were consistently less likely to die of prostate cancer compared to patients with no multiple primary cancers. No studies provided information on comorbidities other than multiple primary cancers.

Race/ethnicity. Fifteen studies (12 SEER or component registries, 1 SEER-Medicare, 1 LAC/USC and 1 using data from the NCHS), covering 1969 to 2007 and reported information on trends in prostate cancer mortality rates (10 studies) or prostate cancer-specific survival (7 studies; 2 studies reported both types of information) stratified by patient race/ethnicity.[16-19,24,26,27,34,40,44,51,60,76,84,86] Eleven of the studies reported exclusively on non-Hispanic whites or blacks whereas four reported on other racial/ethnic groups as well.

Overall, studies demonstrated an increase in the mortality rate from the 1980s to the early-1990s, followed by a decrease from the mid-1990s to 2007 for all racial/ethnic groups. Notably, the mortality rates among blacks were consistently higher compared to that of non-Hispanic whites in all studies and across time periods.

Regarding prostate cancer specific-survival, all studies demonstrated improvements in survival over time, for all racial/ethnic groups during their respective time periods. Six of the seven studies reporting relevant information demonstrated that blacks were at higher risk for prostate cancer death compared to non-Hispanic whites, although the difference between the two groups appeared to decrease over time. One study found no significant difference in the probability of prostate cancer death (versus nonprostate cancer death) between non-Hispanic whites and blacks for the years 1988-95, after adjusting for multiple potential confounders (including tumor grade and stage at diagnosis).[24]

Tumor Characteristics

Stage. Six studies (all SEER based) reported on temporal trends in prostate cancer specific mortality among patients diagnosed with prostate cancer stratified by tumor stage at diagnosis.[18,24,40,78,81,86] Studies covered 1969 to 2001 and demonstrated that over time the proportion of deaths due to prostate cancer among patients diagnosed with the disease has decreased, particularly for patients with early-stage (localized or regional) disease at diagnosis. One study, using the SEER database, demonstrated that the risk of death due to prostate cancer based on tumor stage was persistent from 1988 to 1995, with risk of death more than five times higher for patients diagnosed with distant disease and more than two times higher among those diagnosed with regional disease, as compared to patients with localized disease.[24]

We did not identify any studies reporting trends in population mortality rates stratified by tumor stage.

Tumor grade. Three studies (all SEER based) including prostate cancers diagnosed between 1973 to 1997, reported information on trends in prostate cancer survival stratified by tumor grade at diagnosis.[24,40,86] One of the studies demonstrated that the probability of dying from prostate cancer among patients diagnosed with prostate cancer decreased during the study period (deaths occurring during 1988-95). Although the decrease was observed for all cancer grades, it was more pronounced among patients with well and moderately differentiated tumors.[40] Another study, again using data from the SEER database, compared the probability of death by prostate cancer among patients diagnosed with the disease, stratified by tumor grade. The study demonstrated that, compared to patients with well differentiated tumors, patients with moderately differentiated cancers and poorly differentiated disease had a higher probability of

prostate cancer death (more than two-fold and more than four-fold higher, respectively). These differences were relatively constant over the time period covered by the study (deaths occurring during 1988-95).[24] These findings were confirmed by the third study which, using data from the SEER-Detroit database, showed that patients with grade I tumors have the highest 5-year relative survival compared to those with tumors of higher grade and that improvements in this statistic have occurred for all tumor grades (1 to 4) during the study period (patients diagnosed up to 1997 and followed up to 2001).[86] Trends in mortality rates stratified by tumor grade should be interpreted cautiously given temporal changes in histopathologic grading practice (see next section, subsection on "Histopathologic grading changes").

We did not identify any studies reporting on population trends in mortality rates stratified by tumor grade.

Patient, Tumor, and System-Level Characteristics at Diagnosis

Patient Characteristics

We identified 52 observational studies reporting on patient characteristics at presentation (27 SEER, 7 SEER-Medicare, 10 CaPSURE, 5 NCDB and 3 other databases) (Appendix Tables C1.4-1.14).

The most commonly examined characteristics at baseline were patient age (21 studies),[19,21,29,30,37,42,46,50,52,55,56,63,65,71-75,77,86,91] race/ethnicity (18 studies),[31,37,46,49,52,53,58,61-63,65,71,74,75,77,80,86,92] tumor grade (16 studies),[4,22,30,37,46,52,56,65,67,68,71,73,75,77,86,90] and tumor stage (22 studies).[4,16,23,31,37,46,52,56,62,63,65,67,68,71-76,86,90,91] Information was available for all factors relevant to Key Question 1 except tumor volume. Studies were generally large (median sample = 46,248; 25th-75th percentile, 10,385-134,434), published between 1990 and 2011, and covered the years 1973 to 2008.

Age. Twenty one studies (12 SEER or its component registries, 1 SEER-Medicare, 2 CaPSURE, 5 NCDB, 1 Patterns of Care Study (POCS) data) covering 1973 to 2005, reported information regarding patients' age at presentation (Appendix Table C1.4).[19,21,29,30,37,42,46,50,52,55,56,63,65,71-75,77,86,91]

Seven studies (5 SEER or its component registries, 2 NCDB) reported trends in the average (mean or median) age at diagnosis of prostate cancer. Five of these studies reported reductions in the average age of patients whereas two studies did not report any changes during their respective time periods. Only one of these studies covered the period after 2005: using the SEER database, this study reported a statistically significant reduction over time in the mean age at diagnosis (from 72.2 to 67.2 years, comparing 1988-89 versus 2004-05).[42] This change was statistically significant ($P < 0.001$) and was observed both for whites (absolute reduction = 4.7 years) and blacks (absolute reduction = 6.4 years). Notably, the three studies (all using the SEER database) that reported information on trends in average age stratified by race found that blacks were diagnosed at a younger average age than whites and that this difference persisted over time (i.e., despite changes in the race-specific average age at diagnosis);[19,21,42] no such analyses were reported for individuals belonging to other racial/ethnic groups.

Fifteen studies (7 SEER or its component registries, 1 SEER-Medicare, 2 CaPSURE, 4 NCDB, and 1 POCS) reported the distribution of patients across discrete age groups (one study using the NCDB database provided information on age both as a continuous variable and grouped into discrete categories and is also included in the preceding paragraph) and generally

supported a trend toward younger age at diagnosis (the effect was significant in four of the six studies reporting results of statistical analyses).

Comorbidity. Only two studies (1 CaPSURE, 1 POCS), covering 1997 to 2003, reported trends in the number of comorbidities present at the time of diagnosis of prostate cancer (Appendix Table C1.5).[71,77] The CaPSURE database analysis grouped individuals into three groups: those with no comorbidities, those with one or two comorbidities, and those with three or more comorbidities.[71] The study found no statistically significant difference in the distribution of patients in these groups, when comparing 1997-99 versus 2000-03. The POCS analysis grouped individuals into two groups (those with no comorbidity and those with one or more comorbidities) and compared the frequency of each group across two years (1998 versus 2002).[77] The study concluded that the proportion of patients with no comorbidity has increased over time (from 78.3 percent to 87.4 percent; $P < 0.01$).

Race/ethnicity. Eighteen studies (9 SEER, 3 SEER-Medicare, 2 CaPSURE, 2 NCDB, 2 other databases) covering 1973 to 2003, reported information on trends in the racial/ethnic distribution of patients with prostate cancer (Appendix Table C1.6).[31,37,46,49,52,53,58,61-63,65,71,74,75,77,80,86,92] Seven of the studies analyzed only whites and blacks; the remaining 11 studies considered additional racial/ethnic groups. Generally, there was no consistent pattern in the racial or ethnic distribution of cases over time: some studies indicated that the number of whites increased over time, others that it remained stable, and others that it decreased. Studies using the same database often provided discrepant results even for overlapping time periods; thus, no clear conclusion can be reached. These discrepancies may be due to differences in the selection criteria employed in each study.

Tumor Characteristics

Stage. Twenty two studies (8 SEER, 2 SEER-Medicare, 6 CaPSURE, 5 NCDB, 1 LAC/USC) covering 1973 to 2007, reported information on trends in the distribution of prostate cancer stage at diagnosis (Appendix Table C1.7). [4,16,23,31,37,46,52,56,62,63,65,67,68,71-76,86,90,91] Fifteen of the studies reported information by grouping cases based on information on tumor size, lymph node status, and the presence of distant disease (e.g., by grouping patients into localized, regional, and distant disease stage or by using the American Joint Committee on Cancer staging classification). Six studies reported information on the distribution of T stage groups only (2 for T1-T4, 1 for T1-T3 and 2 for T1-T2a) and one study reported information only on lymph node status.

Studies reporting on cancer stage consistently demonstrated decreases in the proportion of patients presenting distant disease and concomitant increases in the proportion of patients with localized or regional disease, over their respective time periods. All of the studies reporting information on the distribution of T stage used the CaPSURE data. The two studies reporting on T1-T4 tumors[4,65] and the two studies reporting on T1-T3 tumor[71,90]s consistently demonstrated reductions in the proportion of patients presenting with higher T stages (i.e., a shift towards increasing proportion of patients with T1/2 tumors, even among patients classified as “high risk”). The two studies reporting on T1/T2 tumors both demonstrated a decrease of T1a/T1b tumors and T2a tumors and an increase in T1c tumors.[67,68]

The study reporting on lymph node status used the SEER database and suggested that the proportion of patients with positive lymph nodes decreased during the study period (1988-96).[23]

Tumor volume. The large epidemiologic datasets included in our review did not have information pertaining to trends regarding tumor volume. We performed additional targeted searches in MEDLINE using key words relevant to "tumor volume" and time trends, but did not identify additional studies.

Tumor grade. Sixteen studies (7 SEER or its component registries, 6 CaPSURE, 2 NCDB, 1 POCS) covering 1973 to 2007 reported information on trends in tumor grade distribution at disease presentation (Appendix Table C1.8).[4,22,30,37,46,52,56,65,67,68,71,73,75,77,86,90] Studies using SEER data consistently demonstrated reductions in the proportion of patients diagnosed with well- or poorly-differentiated tumors (including undifferentiated tumors) with concomitant increases in the proportion of patients with moderately-differentiated disease. Within each study, this temporal trend was found to be statistically significant in six of the seven studies that reported the results of statistical tests assessing changes in the distribution of tumor grade over time. The six studies using CaPSURE data and one study using POCS data reported temporal trends in the distribution of Gleason scores. Generally, these studies showed a decrease over time in the proportion of tumors with Gleason scores 2-4 and increases in tumors with Gleason scores 5-7.

Prostate specific antigen. Eight studies (7 CaPSURE, 1 POCS), covering 1989 to 2007, reported information on trends in PSA levels at presentation (Appendix Table C1.9).[4,64,65,67,68,71,77,90] Seven studies categorized PSA values (e.g., < 4, 4-10, >10 ng/mL) and only one study reported the median PSA value by diagnosis year. Generally, studies found that the PSA values at diagnosis have decreased over time (i.e., that a larger number of patients are currently diagnosed with PSA concentrations below 10 ng/mL; this was true even for patients with low risk localized prostate cancer[90]).

We did not identify studies reporting on trends in the proportion of screen-detected prostate cancer cases among all cancer cases that met our inclusion criteria. One study demonstrated that for all age groups above 65 years and both for blacks and whites, the proportion of men who underwent PSA testing at least once and were diagnosed with prostate cancer within 90 days of the test among all men undergoing PSA testing has decreased over time (1988-96).[96] Distinguishing between screen-detected prostate cases (i.e., cancer cases identified following investigation triggered by a positive PSA test) and cases where use of the PSA test was used as a confirmatory or add-on test (e.g., as part of the investigation of clinical symptoms or signs suggestive of prostate cancer, or suspicious findings on prostate digital rectal examination) is particularly challenging using administrative data and may be uncertain even after review of complete medical records.[96]

Diagnostic Strategies

Biopsy Frequency. Four studies (2 SEER component registries, 2 SEER-Medicare), covering 1982 to 2001, reported information on trends in the performance of prostate biopsies (Appendix Table C1.10).[52,54,57,89] One study using data from the SEER-Detroit registry reported that the proportion of prostate cancer patients diagnosed through biopsy (compared to those diagnosed through other procedures, such as transurethral resection of the prostate) increased over time (1982-95, $P < 0.001$).[52] A similar trend was evident in a study using data from the SEER-New Mexico registry.[54] A SEER-Medicare study also demonstrated an increase in the age-adjusted rate of biopsy procedures (from 685 to 2600 per 100,000 men) between 1986 and 1991.[57] An updated analysis from the same database, covering the period 1993-2001, reported that the age-

and race-adjusted biopsy rate did not statistically significantly change over the study 9-year period (trend P > 0.2).[89]

Number of cores. A single primary study[70] provided information on the number of biopsy cores obtained during the investigation of suspected prostate cancer cases (Appendix Table C1.11). The study utilized the CaPSURE database and covered 1997 to 2002. It demonstrated a significant increase in the mean number of cores examined per patient (from 7.5 in 1997 to 9.8 in 2002). We note that the study excluded patients who were evaluated with less than six cores, which may have led to underestimation of the change in the number of cores obtained. However, the study found that the increase in the number of cores over time was significant (+0.41 cores per patient per year, P <0.001).

Histopathologic grading changes. In a study of prostate cancer patients from the Connecticut Tumor Registry (1990-92), investigators obtained medical records, pathology reports, and the original slides used for pathological examination for 1858 (49 percent) of the patients diagnosed during the study period.[97] A single pathologist (blinded to the originally assigned Gleason score) regraded all slides (2002-04). The contemporary reading of the slides resulted in the assignment of significantly higher scores compared to the original readings (mean score increase from 5.95 to 6.8; P < 0.001). The study also demonstrated that this reclassification causes an increase in the Gleason-score adjusted prostate-cancer-specific survival; even in the absence of changes in treatment efficacy or tumor biology (since the same patient histories were used such changes cannot explain differences in survival patterns). This observation is often referred to as the "Will-Rogers" [b] phenomenon.[98]

We also identified a structured review on the same topic through additional targeted searches (Appendix Table C1.12).[99] None of the studies included in this review (other than the one discussed above) fulfilled our inclusion criteria.

System Characteristics

Differences in geographical access and other system-level factors. Four studies (2 CaPSURE, 1 POCS, 1 NCDB), covering 1986 to 2003, reported information on changes in the distribution of patients by system-level factors (Appendix Table C1.13).[65,71,74,77] Three studies (2 CaPSURE, 1 POCS) provided information on trends in the distribution of patients' insurance status at diagnosis.[65,71,77] The two studies utilizing CaPSURE data demonstrated a decrease in the proportion of patients with Medicare coverage at the time of diagnosis over the time periods covered (1997-2003 and 1989-2001).[65,71] In contrast, the study using POCS data did not demonstrate a change in the distribution of insurance status over time (1998-2002).[77]

One study, using the CaPSURE database reported on trends in the distribution of settings (community versus academic) and geographic regions over time.[65] Comparing 1997-2001 to 1989-97, there was an increase in the number of patients seen in academic settings (compared to community settings) and an increase in the number of patients originating from Midwestern

[b] The Will-Rogers phenomenon arises when a member who is in the bottom half of a group with a high average outcome is reclassified as a member who is in the top half of a group with a lower average outcome, resulting in increases of the average outcome in both groups. In the prostate cancer case, the phenomenon occurs when the same biopsy sample receives a different Gleason score if assessed using different scoring criteria at different time points (for example, in the 1990s versus now). When members of the low risk group with the least favorable histology are reclassified into the high-risk group (which on average had worse prognosis than the reclassified members), this will inflate grade-adjusted survival over time for both groups.

states (and a concomitant decrease in the proportion of patients from Eastern or Southern states). Because the centers participating in CaPSURE have not remained stable over time, changes in these distributions may be difficult to interpret.

One study, comparing 1998 to 2002, reported an increase in the number of patients residing in areas of higher median income. Again, because sampling strategies changed between the POCS years (and different regions were included), patterns in the distribution of income are difficult to interpret.[77]

Finally, one study, using the NCDB, assessed trends in the distribution of patients by hospital caseload, over time (1986-87 and 1992). There was little evidence of change over the time period covered.[74]

Trends in Treatment Patterns

Twenty one studies (6 SEER, 3 SEER-Medicare, 6 CaPSURE, 5 NCDB, 1 POCS) provided information on treatment trends over time (Appendix Table C1.14).[4,27,36,41,47,48,58,66,67,69,71-75,77,82,87,88,90,91] Studies were generally large (median sample size = 71,602; 25^{th}-75^{th} percentile 6290-142,340), published between 1994 and 2011, and covered 1973 to 2008. In 12 studies, patients managed by observational management strategies of no active treatment (AS, WW or expectant management) were considered in aggregate with patients receiving androgen deprivation therapy (ADT), or it was unclear whether the data allowed the distinction between these treatments. Most studies demonstrated decreasing trends in the proportion of patients being managed with strategies other than surgery or radiotherapy throughout their respective time periods; studies explicitly reporting on AS/WW-type strategies also indicated decreases in the proportion of patients receiving such treatments. In all seven studies (6 CaPSURE, and 1 using POCS data) providing information for years after 2000, the proportion of patients receiving AS/WW was less than 10 percent; this also held true for subgroups of "low-risk disease" (typically defined based on T stage, Gleason score and PSA criteria) investigated in two studies (both using CaPSURE data).

Summary

We reviewed 79 studies based on large epidemiologic databases sourced from the U.S. population. For all age or race/ethnicity groups investigated, the incidence rate appears to have peaked in the early 1990s; subsequently, the incidence rate declined between the early 1990s and 1999. Studies consistently demonstrated that early-stage (localized and regional) prostate cancer cases were responsible for the observed increase in prostate cancer incidence from the mid-1980s up to the mid-1990s. Studies also demonstrated decreases in the prostate cancer-specific mortality rate for all age groups between the early-1990s and 1999. Mean age of diagnosis has decreased over time, both for blacks and whites. Another consistent trend over time has been the decrease in low- and high-grade (corresponding to Gleason scores 2-4 and ≥7, respectively) tumors, and a concomitant increase in intermediate grade tumors (corresponding to Gleason scores 5-6). These trends, and their impact on grade-adjusted mortality statistics, need to be interpreted cautiously in the light of changes in histopathologic grading practice (causing a Gleason score shift toward higher values), and the widespread implementation of PSA screening (which will tend to increase the detection of nonhigh grade tumors). Over time, patients diagnosed with prostate cancer are less likely to die of the disease (i.e., they are more likely to die of nonprostate cancer causes); this is particularly true for patients diagnosed at older age. Most studies demonstrated a decrease over time of the proportion of patients being managed with

strategies other than radical prostatectomy or radiation therapy. Studies explicitly reporting on AS/WW, also indicated decreases over time in the proportion of patients being managed with such observational management strategies; this was true even for subgroups of men with low-risk disease.

There is little doubt that many of the observed trends in the presentation and natural history of prostate cancer in the U.S. in the last 3 decades are at least in part due to the widespread use of PSA screening. However, for many of these trends summarized in the preceding sections there are multiple potential explanations. For example, observed trends in prostate-cancer specific mortality may be explained by the implementation of PSA screening and its impact on ascertaining cause of death[100], improved treatments for localized disease (surgery or radiotherapy), widespread use of ADT, earlier detection (and treatment) of recurrent disease, or changes in the underlying disease biology.[33] Modeling studies may provide insight into the underlying causes of the observed temporal changes[c]; modeling has been used to quantify the impact of PSA testing on population incidence[101], mortality[102,103], and tumor grade at diagnosis[104], or to explore the potential for overdiagnosis.[96,105]

Key Question 2. How are active surveillance and other observational management strategies defined?

There are generally three scenarios in which a man with newly diagnosed prostate cancer might not undergo immediate definitive treatments like RP or RT: 1) his disease has a low risk of rapid progression and therefore it is felt that he could be safely monitored and still receive definitive treatment should the need arise; 2) his disease may have a higher risk of rapid progression but he may not be an ideal candidate for definitive treatments after careful deliberation of the different tradeoffs (e.g., life expectancy gained versus the compromise in quality of life living with side effects from immediate treatments), therefore, he could be followed clinically and be offered palliative treatments should he become symptomatic; or 3) his disease is advanced and only palliative treatments are indicated. In the literature, the first approach (scenario 1) is generally termed "active surveillance (AS)," while the second approach (scenario 2) is generally termed "watchful waiting (WW)." However, it is important to note that investigators have used the terms AS and WW interchangeably. Terms like "expectant management," "conservative management," and others to denote one of the two approaches have also been used. Regardless of the actual term used, we attempt to clarify the intent of the different approaches in summarizing the relevant studies.

AS management strategies typically use a predefined protocol to monitor triggers for initiating curative treatments, whereas watchful waiting (WW) strategies use a somewhat passive (compared to AS) followup and upon symptomatic disease progression palliative treatments are instituted. A wide variety of combinations of monitoring parameters including clinical symptoms, digital rectal examination (DRE) findings, Gleason score, PSA concentrations, PSA doubling time and/or velocity, results from transrectal ultrasound (TRUS) guided rebiopsy, bone scan or other imaging modalities have been used. However, the optimal monitoring strategies in patients choosing AS have not yet been well-characterized.[7]

For this Key Question, we undertook a systematic review of the literature to identify studies that followed men who were initially managed conservatively (e.g., AS and WW) and that

[c] For example, the Cancer Intervention and Surveillance Modeling Network (CISNET), a consortium of NCI-sponsored investigators, uses a comparative modeling approach to explore prostate cancer trends in the US population. Additional information is available at http://cisnet.cancer.gov/; last accessed October 31, 2011.

documented the eligibility criteria for patient selection and followup protocols. We also extracted triggers for recommending treatment and the definitions of prostate cancer progression (Appendix Table C2.1). We considered studies reporting on observational management strategies (i.e., no immediate active treatment with curative intent), enrolling patients based on predefined eligibility criteria, and using prespecified protocols for followup. We reviewed the full-text articles of all qualifying studies but only included the earliest publication or the article with most complete information from the same center or research team that used the same observational strategy (i.e., when the same definition was consistently used in all publications). However, we also considered additional publications reporting on the same cohort when they provided additional relevant information.

There is a broad spectrum of strategies for observational management described in research publications. For the purpose of operationalizing the process of summarizing the various definitions, we divided protocols into those clearly described as having curative intent and those in which their aims were either unclear or primarily palliative, regardless of how these regimens were labeled. This categorization was done for practical reasons, not to suggest what the definitions or protocols for AS, WW, or any other observational strategy should be.

Due to often poor reporting of treatment intent, we based our interpretation on the specific triggers used to initiate treatment in each study. Specifically, we operationally defined protocols with curative intent as those that also used other parameters (with or without symptomatic disease progression) as triggers for active treatment. Protocols where treatment was only initiated upon symptomatic disease progression were considered as palliative.

Protocols With Curative Intent

We identified 16 unique cohorts reporting formal protocols to monitor triggers for curative treatment of prostate cancer. Of these cohorts, ten are in the United States, two in Canada, two in the UK, one in the Netherlands, and one in Japan (Table 1). In all cohorts, AS was offered to men with low-risk or clinically localized prostate cancer although the eligibility criteria varied. The protocols varied across all 16 cohorts. Baylor College of Medicine and Memorial Sloan-Kettering Cancer Center were the first institutions to report enrollment of patients into AS program in 1984.

Table 1. Unique AS cohorts[a]

Cohorts or Centers	Country	Beginning of enrollment (year)
Baylor College of Medicine and Memorial Sloan-Kettering Cancer Center[106]	US	1984
McGill University[107]	Canada	1987
University of Connecticut Health Center[108]	US	1990
Four tertiary care academic medical canters[109]	US	1991
University of Miami[110]	US	1991
University of California at San Francisco[111]	US	After 1991
Royal Marsden Hospital[b112]	UK	1993
Johns Hopkins[113]	US	1994
Toronto- Sunnybrook Regional Cancer Center[114]	Canada	1995
Memorial Sloan-Kettering Cancer Center[115]	US	1997
ProtecT (Prostate testing for cancer and Treatment Trial)[116]	UK	2000
Dana-Farber Cancer Institute[117]	US	2000
Kagawa Medical University[c118]	Japan	2002
Cleveland clinic[119]	US	2004
PRIAS (Prostate cancer Research International Active Surveillance)[120]	Netherlands	2006
PASS (the Canary Prostate Active Surveillance Study)[121]	US	2008

[a] Cohorts are listed chronologically by enrollment start year. Some cohorts had multiple publications providing different pieces of information on eligibility criteria and followup protocol. In this case, only the earliest publication was used as the primary citation of the cohort. Details are described in Tables 3 and 5.

[b] Royal Marsden Hospital had both AS and WW protocols (described separately). Since 1993, the Royal Marsden Urology Unit has offered an AS policy as a management option for favorable-risk early prostate cancer.

[c] The cohort was not an AS cohort before 2002. See the next section, "Observational management strategies with palliative intent," for its earlier eligibility criteria and followup protocol.

Common Metrics: Eligibility Criteria for Low-risk or Clinically Localized Prostate Cancer in AS Cohorts (Table 2-3)

The criteria used to identify patients with low-risk or clinically localized prostate cancers generally varied across the 16 cohorts, with the exception that no cohorts enrolled patients with clinical stage greater than T2. The most commonly used parameters of patient eligibility criteria for AS were Gleason score (12 cohorts), PSA (10 cohorts) and number of biopsy cores positive for cancer (8 cohorts).

Age. Three cohorts reported using age as part of their patient eligibility criteria for AS.[109,118,122] The age criterion was less than 75 years in one multicenter cohort (Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami),[109] less than 80 years in the cohort at University of Miami,[122] and between 50 and 80 years old in the cohort at Kagawa Medical University in Japan.[118] One cohort justified the use of age as part of patient inclusion criteria to mirror those patients who would otherwise be eligible for RP or RT due to a life expectancy greater than 10 years at the time of diagnosis;[109] the other two did not report the reason for including age as part of patient eligibility criteria for AS.[118,122]

Gleason score. Twelve cohorts (16 publications) used the Gleason score as part of patient eligibility criteria for AS (Tables 2 and 3).[106,109-115,117-120,122-125] Nine of the 12 cohorts used Gleason score 6 or less (or no pattern 4 or 5). The remaining three cohorts allowed Gleason pattern 4.[111,112,114] Toronto-Sunnybrook Regional Cancer Center cohort changed the Gleason score criterion that was used to define "favorable-risk" patients for offering AS in the beginning of the study due to the publication of more convincing evidence of a significant difference in natural history between Gleason 6 and 7.[126] Specifically, between 1995 and 1999, AS was offered to all patients who had a Gleason score 6 or less and PSA 10 ng/mL or less, in addition to

older patients (age ≥70 years) with Gleason up to 3+4 or PSA up to 15 ng/mL. Since January 2000, the cohort was restricted to patients who had a Gleason score 6 or less and PSA 10 ng/mL or less, regardless of age.[114,125]

Number of cores positive for cancer. Eight cohorts (11 publications) used a maximal number of biopsy cores positive for cancer as part of patient eligibility criteria for AS.[109,110,112,113,115,117,118,120,122-124] Two criteria were used: 2 or fewer (5 cohorts) and 3 or fewer (3 cohorts) positive biopsy cores. It should be noted that the biopsy strategies varied across these cohorts. For example, some cohorts used sextant (6-core) biopsy, some used octant (8-core) biopsy, and others performed extended biopsy (>10 cores).

Percentage cancer involvement in each core. Five cohorts (7 publications) used "low-volume disease" as part of patient eligibility criteria for AS.[110,111,113,115,117,122,123] In three cohorts, the definition of "low-volume disease" was involvement of less than half of any individual core with cancer. In the other two cohorts (3 publications), the criterion percent of biopsy cores with cancer involvement was described variably as less than half of two biopsy cores,[110] less than 20 percent in one or two biopsy cores,[123] and cancer involvement of less than 33 percent of biopsy cores.[111]

PSA. Ten cohorts (14 publications) used PSA as part of patient eligibility criteria for AS (Table 2.2).[109-115,118-120,122-125] Three PSA thresholds were used: ≤10 (7 cohorts), ≤15 (3 cohorts), and ≤20 (2 cohorts) ng/mL. Two cohorts used two different PSA density thresholds (≤0.15 or ≤ 0.2 ng/mL/cm^3). The number of cohorts do not sum to the total number of cohorts because some cohorts used multiple criteria. Of these, two cohorts reported changes in the PSA criteria to a lower threshold in the more recent years. The cohort at Royal Marsden Hospital changed the PSA threshold from less than or equal to 20 ng/mL to less than or equal to 15 ng/mL in 2002,[112,124] and the cohort at University of Miami changed the PSA threshold from less than or equal to15 ng/mL to less than or equal to 10 ng/mL in more recent publications (Tables 2 and 3).[110,122,123]

Table 2. Cohorts that used Gleason score and PSA as part of patient eligibility criteria for AS

Cohorts or Centers	Gleason score	PSA (ng/mL)
Baylor College of Medicine and MSKCC[106]	<7	
McGill University[107]		
UCHC[108]		
Four tertiary care academic medical canters[109]	⩽6	⩽10
University of Miami[110]	⩽6	⩽15; then changed to ⩽10
UCSF[111]	⩽6 with *a priori* exceptions[b]	<10 with *a priori* exceptions[d]
Royal Marsden Hospital[112]	<3+4	⩽20 ; then changed to ⩽15
Johns Hopkins[113]	⩽6	PSA-d ⩽0.15
Toronto[114]	⩽6; ⩽3+4 (if ⩾70 yr)[c]	⩽10; <15 (if ⩾70 yr)[c]
MSKCC[115]	No Gleason grade 4 or 5	<10
ProtecT (Prostate testing for cancer and Treatment Trial)[116]		
Dana-Farber Cancer Institute[117]	⩽6 with no pattern 4	
Kagawa Medical University[118]	⩽6	⩽20
Cleveland clinic[119]	No Gleason grade 4 or 5	⩽10
PRIAS[120]	⩽3+3=6	⩽10; PSA-d ⩽ 0.2
PASS[121]		

Grey shading indicates that a study did not report use of a particular criterion; PSA-d = PSA density (ng/mL/cm3) which is calculated by dividing the PSA level by the total volume of the prostate.

[a] The Gleason patterns range from 1 to 5, being 5 the least differentiated pattern. A Gleason score of 4+3 (primary + secondary grade) = 7 is different from 3+4 = 7, and has a different prognosis as well. It is therefore more informative to give both patterns, than just providing the sum of those (i.e., the Gleason score).

[b] Exception: having a Gleason pattern of 4 reported only in a microfocus of tumor

[c] (if ≥if yr) = for older patients (age ≥yr years), less than Gleason up to 3+4 criterion and PSA less than 15 was used.

[d] Exception: <15 for men with concurrent benign prostatic hyperplasia or prostatitis

Imaging. One cohort reported that a chest radiograph was mandatory, and bone scan and computed tomography scan of the abdomen and pelvis were performed at the clinician's discretion.[114,125] One other cohort reported that magnetic resonance imaging of the prostate was selectively used at diagnosis.[109]

Behavioral indicators. No behavioral indicator was used explicitly as a criterion for AS program enrollment.

Table 3. Eligibility criteria for enrollment in protocols with curative intent in chronological order of starting enrollment year

Center, Country [Pubmed ID] Enrollment years	Term used in original article	Age (yr)	Gleason score	# biopsy cores /% cores	PSA (ng/mL)	Imaging	Stage	Behavioral indication (other than patients' choice or preference)	
Baylor College of Medicine and MSKCC, US[106] [15017211] 1984-2001	EM /deferred therapy	–	<7	TRUS guided sextant biopsy	–	–	–	Decision for deferred therapy was made by the patient and treating physician together based on the likely presence of small volume cancer.	
McGill Univ., Canada[107] [18484590] 1987-2002	WW; AS	–	–	TRUS guided biopsy	–	–	"Clinically localized caner"[26])	Limited life expectancy because of advanced age or poor medical condition	
Univ. of Connecticut Health Center, US[108] [18707696] 1990-2006	AS	–	–	–	–	–	"low-risk disease"	Patients who elected WW or AS program. Men on WW were generally older with localized prostate cancer who did not desire aggressive intervention. Men on AS were generally younger with low-risk disease.	
Four tertiary care academic medical canters,[a] US[109] [19233410] 1991-2007	AS	⩽75	⩽6	⩽3 positive cores at diagnostic biopsy	⩽10	MRI of the prostate was selectively used at diagnosis	T1-T2a	–	
Univ. of Miami, US[110,122,123] [17850361; 20800964; 21215429] 1991-2007	AS; WW[123]	⩽80[122]	⩽6[110]	TRUS guided biopsy ⩽50% of 2 biopsy cores[110] ⩽2 biopsy cores with ⩽20% in each core[123]	⩽15[110] ⩽10[122,123]	–	⩽T2/T2b[110]	MD influence had the greatest impact on choosing AS (73%), concerns for incontinence (48%) and erectile dysfunction (44%) also reasons for choosing AS[122]	

Table 3. Eligibility criteria for enrollment in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Term used in original article		Age (yr)	Gleason score	# biopsy cores /% cores	PSA (ng/mL)	Imaging	Stage	Behavioral indication (other than patients' choice or preference)	
UCSF, US[111,133] [18433013] >1991	AS	–	⩽6; no Gleason pattern 4 or 5 A priori exception: having a Gleason score of 4 reported only in a microfocus of tumor[133]	Cancer involvement of <33% of biopsy cores <50% of the length of a tumor core involved by carcinoma[133] A priori exception: having >33% positive biopsy cores due to a tumor microfocus[133]		<10 A priori exception: <15 for men with concurrent benign prostatic hyperplasia or prostatitis[133]	–		T1/T2a	life expectancy >3 mo[133]
Royal Marsden Hospital, UK[112,124,134] [15839912; 17850368] 1993-2002; ⩾2002;[124] 2004-2006[134]	AS	–	⩽7 (primary ⩽3) <3+4[134]	Less than half of the biopsy cores positive (octant biopsy).[124] <20% core positive[134]		⩽20[112] <15[124]	–		T1–2, N0/X, M0/X	–
Johns Hopkins, US[113] [20439642] 1994-2008	AS (or EM with curative intent)	–	⩽6	⩽2 cores cancer positive; ⩽50% cancer in any single core		PSA density (PSA before diagnosis divided by prostate volume) ⩽0.15 ng/mL/cm3	TRUS to determine PSA density		T1c	–
Toronto-SRCC, Canada[114,125] [11395227; 19917860] 1995-2002 as a phase II trial; 2003-ongoing as an open prospective cohort	WW; AS[b]	–	⩽7[114] ⩽6; ⩽3+4 (if ⩾70 yr)[125]	–		⩽15[114] ⩽10; <15 (if ⩾70 yr)[125]	Chest X-ray, TRUS of the prostate were mandatory. Bone scan and CT scan of the abdomen and pelvis were performed at the clinicians' discretion.		T1b-T2b N0 M0 (1997 TNM classification)[c]	–

Table 3. Eligibility criteria for enrollment in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Term used in original article	Age (yr)	Gleason score	# biopsy cores /% cores	PSA (ng/mL)	Imaging	Stage	Behavioral indication (other than patients' choice or preference)	
Memorial Sloan-Kettering Cancer Center, US[115] [21167529] 1997-2009	AS	–	No Gleason grade 4 or 5	⩽3 positive biopsy cores (minimum 10), no biopsy core containing >50% cancer involvement	<10	–		T1-T2a	–
ProtecT, UK[116] [19603015] 2000-2008	Active monitoring	–	–	–	–	–		"Clinically localized prostate cancer"	–
Dana-Farber Cancer Institute, US[117] [21167525] 2000-2010	AS	–	⩽6 with no pattern 4	<3 cores positive for cancer and < 50% of cancer in any core	–	–		T1c-T2c	–
Kagawa Medical Univ., Japan[118] [18272471] 2002-2003	AS	50-80	⩽6	TRUS-guided six sextant biopsy 1-2 positive cores per 6-12 systematic biopsy cores	⩽20	–		T1cN0M0	–
Cleveland clinic, US[119] [21256549] 2004-2009	Surveillance	–	No primary or secondary Gleason scores 4 or 5	–	⩽10 (part of D'Amico criteria)	–		Clinical stage T2a or fewer (part of D'Amico criteria)	–
PRIAS, Netherlands[120] [19817747] 2006 – ongoing	AS	–	⩽3+3=6	TRUS guided biopsy[d] Adequate biopsy sampling according to biopsy protocol; maximal 2 biopsy cores invaded with prostate cancer	⩽10 PSA density ⩽ 0.2 ng/ml/ml	–		T1c or T2	–

Table 3. Eligibility criteria for enrollment in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years		Term used in original article	Age (yr)	Gleason score	# biopsy cores /% cores	PSA (ng/mL)	Imaging	Stage	Behavioral indication (other than patients' choice or preference)	
PASS, US[121] [19758683] 2008 – ongoing[e]	AS	–	–	If diagnosis ≤1 year of entry, at least 1 biopsy with ≥10 cores. If diagnosis >1 year of entry, minimum of 2 biopsies; 1 ≤2 years before entry.	–		–		Clinical stage T1-2, NX/0, MX/0	–

DT = doubling time; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; TURP = transurethral resection of the prostate; yr = yr(s); wk = wk(s); mo = mo(s); SRCC = Sunnybrook Regional Cancer Center; BCCA = the British Columbia Cancer Agency; DRE = digital rectal examination; WW = watchful waiting; AS = active surveillance; EM = expectant management; PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = Transrectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; SRCC = Sunnybrook Regional Cancer Center; BCCA = the British Columbia Cancer Agency; ED = erectile dysfunction; PRIAS = Prostate cancer Research International Active Surveillance; ProtecT = Prostate testing for cancer and Treatment; UCSF=University of California at San Francisco; European Randomized Study of Screening for Prostate Cancer = ERSPC; VA = Veterans Affairs; MSKCC = Memorial Sloan-Kettering Cancer Center
√ = item was used as part of monitoring strategy but explicit criteria were not defined
– = item was not used or not reported as part of monitoring strategy

[a] Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami
[b] Multiple terms used across multiple publications and within single publication
[c] Since January 2000, Toronto-SRCC study was restricted to low-risk patients only.
[d] PRIAS protocol can be found: http://www.erspc-media.org/media/publications/PRIAS%20Project_background.pdf. Accessed September 28, 2011.
[e] Prostate Active Surveillance Study (PASS) protocol can also be found at Clinicaltrial.gov (NCT00756665): http://clinicaltrials.gov/ct2/show/NCT00756665. Accessed September 28, 2011.

Followup Protocols of AS Cohorts (Tables 4 and 5)

All 16 cohorts included regular PSA testing in the followup protocol (Tables 4 and 5). DRE was included in the followup protocols of 12 of the cohorts. Fourteen cohorts included routine rebiopsy (Figure 5). The testing frequency of PSA, DRE, and rebiopsy varied across the cohorts. One cohort also incorporated a regular bone scan schedule. Criteria for recommending curative treatments varied across the cohorts. The recommended treatments were not standardized and were left at the discretion of treating physicians and patients in many of the cohorts.

Gleason score. Twelve cohorts (21 publications) used the Gleason score as part of monitoring criteria for disease progression.[106-108,110-115,117,119-121,124,125,127-132] Generally, disease progression was defined as a Gleason score or pattern greater than those used in the eligibility criteria for AS (Table 2). One cohort did not use the Gleason score as part of eligibility criteria for AS, but reported the use of any increase in the Gleason grade as part of monitoring criteria for disease progression.[121]

Number of cores positive for cancer. Eight cohorts (9 publications) used the minimal number of biopsy cores positive for cancer as part of monitoring criteria for disease progression.[106-108,110,113,115,117,118,120] Two criteria were used: 3 or more (6 cohorts) and greater than 4 (3 cohorts) positive biopsy cores. One cohort reported that "an increased number of cores positive for cancer" was used as one of the parameters for defining disease progression but the specific number of cores was not reported.[108] The rebiopsy frequencies varied across the cohorts (Figure 5).

Percentage cancer involvement in each core. Six cohorts (8 publications) used more than 50 percent cancer involvement in each biopsy core as part of monitoring criteria for disease progression.[107,112,113,115,117,124,128,132] Two other cohorts considered an increase in tumor volume as part of monitoring criteria for disease progression, but specific percentage cancer involvement was not reported.[108,119]

PSA. All 16 cohorts included regular PSA testing in the followup protocol. The testing frequency varied across the cohorts (Figure 5). Six cohorts considered rising PSA and/or PSA kinetics as part of triggers for treatment but did not specify the detailed criteria. One cohort explicitly reported that PSA kinetics was not used as part of triggers for intervention. Nine cohorts used a variety of PSA triggers for treatment (Figure 5): PSA doubling time ranging from <1 to <4 years (6 cohorts), various PSA velocity criteria (4 cohorts), and PSA greater than 10 ng/mL (1 cohort). The number of cohorts do not sum to the total number of cohorts because some cohorts used multiple criteria. Of these, Toronto-Sunnybrook Regional Cancer Center cohort changed the original PSA doubling time trigger (PSA doubling time < 2 years in the first 4 years of the study) to PSA doubling time < 3 years in 1999.[130] In 2005, the same group also added risk zone to the protocol (the group developed a clinical decision making aid that can define 3 risk zones of high, intermediate and low risk of reclassification when overlaid with PSA data from each patient).[130] A patient with a PSA consistently in the high risk zone is recommended to undergo treatment. Two cohorts did not use PSA kinetics as a trigger for treatment.[113,117]

Imaging. One cohort performed annual bone scan for the first 2 years and biennially thereafter.[114,125,129,130] Another cohort reported that magnetic resonance imaging of the prostate was selectively performed every 1 to 3 years during followup.[109]

Behavioral indicators. No study based used a formal assessment of any behavioral indicator to initiate active treatment as part of their followup protocol, but one cohort explicitly reported that some patients requested treatment due to anxiety related to increasing PSA.[108]

Figure 5. Followup frequencies of 16 unique AS cohorts[a]

AS cohort	Year	1				2				3				4			
	Month	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48
Baylor College of Medicine and MSKCC[106]	PSA-test	√	√	√	√		√		√		√		√		√		√
	DRE	√	√	√	√		√		√		√		√		√		√
	Rebiopsy		√														
McGill University[107]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy				√				√				√				√
UCHC[108]	PSA-test		√		√		√		√		√		√		√		√
	DRE				√				√				√			√	
	Rebiopsy								√								
Four academic medical canters[109]	PSA-test				√				√				√			√	
	DRE				√				√				√			√	
	Rebiopsy	within 18 mo; then every 1 to 3 yr															
University of Miami[110]	PSA-test	√ *	√ *	√ *	√ *	√ *	√ *	√ *	√ *		√		√		√		√
	DRE	√ *	√ *	√ *	√ *	√ *	√ *	√ *	√ *		√		√		√		√
	Rebiopsy	within 9-12 mo							√				√				√
UCSF[111]	PSA-test	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Rebiopsy					√								√			
Royal Marsden Hospital[112]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy	not routine															
Johns Hopkins[113]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy				√				√				√				√
Toronto[114]	PSA-test	√	√	√	√	√	√	√	√		√		√		√		√
	DRE																
	Rebiopsy									√ **							
	Bone scan				√				√								√
MSKCC[115]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy	within 12-18 mo; then every 2 to 3 yr															
ProtecT[116]	PSA-test	√	√	√	√		√		√		√		√		√		√
	DRE																
	Rebiopsy	not routine															
Dana-Farber Cancer Institute[117]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy					√				√				√			

Figure 5. Followup frequencies of 16 unique AS cohorts[a] (continued)

AS cohort	Year	1				2				3				4			
	Month	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48
Kagawa Medical University[118]	PSA-test	√** *		√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE																
	Rebiopsy				√												
Cleveland clinic[119]	PSA-test				√				√				√			√	
	DRE																
	Rebiopsy								√								√
PRIAS[120]	PSA-test	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE		√		√		√		√				√				√
	Rebiopsy				√												√
PASS[121]	PSA-test	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE		√		√		√		√				√				√
	Rebiopsy				√				√								√

Legends: √* = every 3-4 mo; √** = every 3-4 years; √*** = every 2 mo for 6 mo. Merged cell represents a range of followup frequency; for example, a merged cell of 3 and 6 mo with a check mark in the middle of the merged cell means a followup frequency of 3 to 6 mo.

MSKCC= Memorial Sloan-Kettering Cancer Center; UCHC= University of Connecticut Health Center; Four academic medical centers=Cleveland Clinic Foundation, University of British Columbia and University of Miami; UCSF= University of California at San Francisco; ProtecT= Prostate testing for cancer and Treatment Trial; PRIAS= Prostate cancer Research International Active Surveillance; PASS= the Canary Prostate Active Surveillance Study)

[a] Cohorts of active surveillance in chronological order of starting enrollment year. Some cohorts had multiple publications providing different pieces of information on followup protocol. In this case, only the earliest publication was used as the primary citation of the cohort. See Table 2.5 for more detailed monitoring criteria in each cohort.

Table 4. Cohorts that used PSA (ng/mL) or PSA kinetics as part of followup protocol for AS

Cohorts or Centers	PSA (ng/mL)
Baylor College of Medicine and MSKCC[106]	PSA velocity >0.75 ng/mL/year
McGill University[107]	Used but not specified
UCHC[108]	Used but not specified
Four tertiary care academic medical canters[109]	Used but not specified
University of Miami[110]	PSA increase 25-50 %/year
UCSF[111]	PSA velocity >0.75 ng/mL/yr; PSA DT <1 year
Royal Marsden Hospital[112]	PSA DT<4 years; PSA velocity >1 ng/mL/year
Johns Hopkins[113]	PSA kinetics were not used as part of triggers for intervention
Toronto[114]	PSA DT <2 years
MSKCC[115]	>10
ProtecT (Prostate testing for cancer and Treatment Trial)[116]	Used but not specified
Dana-Farber Cancer Institute[117]	Used but not specified
Kagawa Medical University†[118]	PSA DT <2 years
Cleveland clinic[119]	Used but not specified
PRIAS[120]	PSA DT 0 to 3 years
PASS[121]	PSA DT <3 years

PSA DT = PSA doubling time which is defined as the time PSA needs to double its start-value. PSA V = PSA velocity (ng/mL/yr) which is the absolute increase of PSA values in one year.

Table 5. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Baylor College of Medicine and MSKCC, US[106] [15017211] 1984-2001	DRE and PSA every 3 mo first yr and every 6 mo thereafter.	any new Gleason pattern 4 or 5	Repeat TRUS guided sextant biopsy was recommended at 6 mo: bilateral or multifocal cancer, or > 4 cores with cancer	PSA velocity was calculated from 3 separate recorded values in a 12-mo period: > 0.75 ng/ml/yr in 12 mo, or 24 mo	–	–	–	Definitive treatment when objective progression or patients' requests.
McGill Univ., Canada[26,107] [18484590] 1987-2002	Every 3-6 mo PSA and DRE	Gleason pattern of 4	TRUS guided biopsy was done annually or when there was a change in DRE or PSA. ⩾3 positive, or >50% cancer in at least 1 core	√	–	√	–	Clinical disease progression on DRE or repeated sextant biopsy, patient preference, or rising PSA level.[26]
Univ. of Connecticut Health Center, US[108] [18707696] 1990-2006	Every 3-6 mo PSA, DRE every 6 to 12 mo, rebiopsies recommended 2 yr after initial biopsy	Progression in Gleason score	Increase in tumor volume (increased number or percent of cores positive)	√	–	Anxiety related to increasing PSA trend	–	Increase in tumor volume, progression in Gleason score, onset of urinary symptoms, change in DRE or patient request (due to anxiety related to increasing PSA trend).

Table 5. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests		Triggers for interventions
Four tertiary care academic medical canters,[f] US[109] [19233410] 1991-2007	Every 6-12 mo PSA and DRE, rebiopsies within 18 mo and then every 1 to 3 yr	–	–		√	MRI of the prostate was selectively every 1 to 3 yr	–	–	Criteria for recommending treatment were nonstandardized and physician specific.
UCSF, US[111,131-133,135] [18433013; 21115873; 21419438] >1991	Every 3 mo PSA and DRE; prostate biopsy every 12-24 mo (after 2003)	Gleason upgrade to ⩾4 (if (⩽6 at diagnosis) or ⩾4+3 (if 3+4 at diagnosis)[131]	TRUS guided biopsy every 6-12 mo ⩾33% of cores or >50% of any core[132] Rebiopsy ≥10 cores[135]		PSA velocity >0.75 ng/mL/yr PSA DT < 1 yr[133]	–	–	–	Disease progression; no specific protocol for intervention (implied)
Univ. of Miami, US[110,127] [17850361; 10759669] 1991-2007	Every 3-4 mo PSA and DRE for 2 yr and every 6 mo thereafter.[110]	⩾7[110]	>2 positive cores (After 2000, a peripherally targeted TRUS biopsy of 10-12 cores was performed 9-12 mo after the first rebiopsy, and then annually or earlier if dramatic rise in PSA or a change on DRE.)[110]		Biochemical progression: PSA increase 25-50 %/yr[127]	TRUS (needed for determining tumor volume)[110]	–	–	Treatment is encouraged at an increase in tumor volume, Gleason score ⩾7, or the presence of >2 positive cores at rebiopsy.[110] Treatments were offered at the time of local stage progression by DRE and/or biochemical progression, or systemic progression.[127]

Table 5. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests		Triggers for interventions
Royal Marsden Hospital, UK[112,124,128] [15839912; 17850368; 18949747] 1993-2002; ⩾2002;[124] 2004-2006[134]	Every 3-6 mo PSA and DRE for 2 yr and every 6 mo thereafter. Rebiopsy not routine.[112] After 2002; monthly PSA in yr 1, every 3 mo in yr 2, and every 6 mo thereafter.[124,134] DRE every 3 mo for 2 yr.	>7[124] Primary Gleason ⩾4, (initial Gleason 3+3, upgraded to Gleason ⩾3+4)[134,136]	TRUS-guided octant biopsy at 18-24 mo. Sextant or octant ⩾50% biopsy cores positive.[124]	PSA DT<4 yr[124] PSA velocity >1 ng/mL/yr based on a minimum of 4 values observed over minimum of 6 mo[128,134]		Repeat imaging only if clinically indicated.[112]	–	–	Rate of rise of PSA, according to judgment of each patient and clinician.[112] PSA DT<4 yr, histologic progression, or patient preference, or PSA velocity >1 ng/mL/yr [128]
Johns Hopkins, US[113] [20439642] 1994-2008	Every 6 mo PSA and DRE; annual extended 12-core biopsy	⩾7; or Gleason pattern 4 or 5	>2 cores cancer positive; or single core >50% cancer (from annual extended 12-core biopsy)	√ (PSA kinetics not used as a trigger for intervention)		–	Patient request for curative treatment[137]	–	Annual surveillance biopsy: Gleason ⩾7; or Gleason pattern 4 or 5; or >2 cores cancer positive; or single core >50% cancer.

Table 5. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests		Triggers for interventions
Toronto-SRCC, Canada[114,125,129,130] [11395227; 19917860; 20478589; 20846681] 1995-2002 as a phase II trial; 2003-ongoing as an open prospective cohort	Every 3 mo for the first 2 yr and every 6 mo thereafter	Histologic progression: Gleason score upgraded to ⩾8 in the rebiopsy of the prostate at 18 months post enrollment	Subsequent biopsies were performed every 3-4 yr to identify biologic progression.[125] Sextant biopsies were used from 1995 to 2000; since 2000, 10 to 14-core biopsies were performed using the Vienna nomogram.[129]	PSA progression: PSA DT <2 yr, based on at least 3 separate measurements over a minimum of 6 mo; final PSA >8 ng/ml; p-value <0.05 from regression of ln(PSA) on time. Protocol changes in PSA DT assessment or calculation in 1999 and after 2002.[g] In 2005 the group developed a general linear mixed model as a clinical decision making aid.[h130]	Bone scan annually for the first 2 yr and biennially thereafter. If PSA >15 ng/ml, annual bone scan was performed. TRUS was performed every 6 mo.	–	PAP and serum creatinine	Clinical,[i] histologic or PSA progression triggered the offer of treatment based on age, extent of disease and comorbidities.
Memorial Sloan-Kettering Cancer Center, US[115] [21167529] 1997-2009	Every 6 mo PSA and DRE; biopsy was within 12 to 18 mo starting AS and repeated every 2 to 3 yr	Gleason grade 4 or 5	>3 positive biopsy cores (minimum 10), biopsy core containing >50% cancer involvement	>10	–	–	–		Treatment was recommended when the patient no longer met study eligibility criteria during followup.

Table 5. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests		Triggers for interventions
ProtecT, UK[116] [19603015] 2000-2008	PSA every 3 mo in year 1 and ever 6 mo thereafter	–	referred to biopsy if a PSA ⩾3 ng/mL; rebiopsy was not routine	√	–	–	–	The aim was "to identify developing cancers early enough to allow treatment with surgery or radiotherapy"[j] (implied using PSA level or change and/or rebiospy results as triggers)
Dana-Farber Cancer Institute, US[117] [21167525] 2000-2010	PSA and DRE every 6 mo; biopsy every 12 to 18 mo	⩾7	20-core biopsy; ⩾3 positive cores, or >50% of any core involved with cancer	√	–	–	–	Patients with progression[k] were offered surgery or radiotherapy.
Kagawa Medical Univ., Japan[118] [18272471] 2002-2003	PSA every 2 mo for 6 mo, every 3 mo thereafter; Re-biopsy at 1 yr (no data beyond 1 yr)	–	Rebiopsy did not fit initial pathology criteria (i.e., 1-2 positive cores per 6-12 systematic biopsy cores)	PSA DT <2 yr after 6 mo (based on all PSA or most recent 1 yr)	–	–	–	PSA DT <2 yr after 6 mo; rebiopsy did not fit initial pathology criteria
Cleveland clinic, US[119] [21256549] 2004-2009	PSA every 6-12 mo, surveillance biopsy usually every 2 yr or sooner	√	√	√	–	–	–	Considering multiple parameters (PSA and PSA kinetics, changes in DRE, quantity of cancer in biopsy specimens, and biopsy Gleason score)

Table 5. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests		Triggers for interventions
PRIAS, Netherlands[120] [19817747] 2006 – ongoing	PSA at 3 mo, DRE at 6 mo and standard rebiopsy after 1 yr	>3+3=6	Biopsy protocol[l]	PSA DT 0 to 3 yr		–	–	–	PSA DT 0 to 3 yr, T state >2 or rebiopsy findings exceed study inclusion thresholds
PASS, US[121] [19758683] 2008 – ongoing	PSA every 3 mo, DRE every 6 mo, rebiopsy at 6-12 mo from the time of entry, at year 2, then every 2 yr	Any increase in Gleason grade	–	PSA DT <3 yr		–	–	–	Biochemical progression; clinical progression (a stepwise increase in tumor stage by DRE or identification of regional or distant metastasis)

DT = doubling time; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; TURP = transurethral resection of the prostate; yr = yr(s); wk = wk(s); mo = mo(s); SRCC = Sunnybrook Regional Cancer Center; BCCA = the British Columbia Cancer Agency; DRE = digital rectal examination; WW = watchful waiting; AS = active surveillance; EM = expectant management; PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = Transrectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; SRCC = Sunnybrook Regional Cancer Center; BCCA = the British Columbia Cancer Agency; ED = erectile dysfunction; PRIAS = Prostate cancer Research International; ProtecT = Prostate testing for cancer and Treatment; UCSF=University of California at San Francisco; European Randomized Study of Screening for Prostate Cancer = ERSPC; VA = Veterans Affairs; MSKCC = Memorial Sloan-Kettering Cancer Center

√ = item was used as part of monitoring strategy but explicit criteria were not defined

– = item was not used or not reported as part of monitoring strategy

[f] Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami

[g] For the first 4 yr of the study, PSA DT <2y was used as a trigger. This criterion identified 10% of patients as high-risk and was considered overly stringent. In 1999 the cut-off was increased to 3 yr. From 1995 to 2002 PSA DT was calculated by a statistician using linear regression of all PSA values after the patient left the clinic and the 95% upper bound confidence limit of PSA DT had to be <3 yr. Later PSA DT was calculated by physicians who used PSA fluctuations to determine whether PSA DT was "truly" <3 yr.

[h] The model generates 2 reclassification curves (high and low risk) which, when overlaid over PSA data of each patient, defines 3 risk zones of high, intermediate and low risk of reclassification. A patient with a PSA consistently in the high risk zone is recommended to undergo treatment.

[i] Clinical progression = at least one of the following: >2 times of the product of the maximum perpendicular diameters of the primary lesion as measured digitally; symptoms requiring TURP; development of ureteric obstruction; radiological or clinical evidence of distant metastasis.

[j] Source: http://www.epi.bris.ac.uk/protect/

[k] Progression criteria: (1) 3 or more positive cores, (2) increased grade (Gleason score 7 or greater) and/or (3) more than 50% of any core involved with cancer.

[l] PRIAS protocol can be found: http://www.erspc-media.org/media/publications/PRIAS%20Project_background.pdf (assessed 7/15/2011)

Observational Management Strategies with Palliative Intent

We identified 13 unique cohorts reporting followup protocols for patients who initially received no treatment and who were subsequently treated only for symptomatic progression. We labeled these observational management strategies as having primarily palliative intent. Of these cohorts, seven are in the United States; two in Canada; four in the UK; one in Sweden; one across Finland, Sweden, and Iceland; one in the Netherlands; and one in Taiwan (Table 6). Six cohorts were formed in the pre-PSA screening era. Howard University College of Medicine was the first institution to report enrollment of patients into an observational management program in 1967.

Table 6. Unique 13 cohorts of observational management strategies with palliative intent

Cohorts or Centers	Country	Beginning of enrollment (year)
Howard University college of Medicine[138]	US	1967[a]
Orebro Medical Center[139]	Sweden	1977 [a]
North Stockholm[140]	Sweden	1978 [a]
Freeman Hospital[141]	UK	1978 [a]
Western General Hospital[142]	UK	1978 [a]
Taichung Veterans Hospital[143]	Taiwan	1983 [a]
Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4) trial[144]	Finland, Sweden, and Iceland	1989
Erasmus University Hospital[145]	Netherlands	Before 1990
Royal Marsden Hospital (before 2002)[112]	UK	1993
Prostate Cancer Intervention Versus Observation Trial (PIVOT)[146]	US	1994
Watchful Waiting Study[147]	US	1998
Hospitals in Manchester region[148]	UK	Not reported (publication year 2001)
University of Florida[149]	US	2003

Hospitals in Manchester region= University Hospital of South Manchester, Withington Hospital, Christie Hospital; Hope Hospital. Some cohorts had multiple publications providing different pieces of information on eligibility criteria and followup protocol. In this case, only the earliest publication was used as the primary citation of the cohort. Details are described in Table 7 and 8.

[a] Early cohorts that did not use PSA as part of eligibility criteria or followup protocols. These cohorts are assumed to have formed during the pre-PSA screening era.

Common Metrics: Eligibility Criteria for Observational Management Strategies with Palliative Intent (Table 7)

The six cohorts in the pre-PSA screening era enrolled patients primarily based on clinical staging alone. Of these, four cohorts enrolled patients with clinical stage T2 or less,[138,139,142,143] one cohort enrolled "patients without symptoms after initial outflow tract surgery or biopsy,"[141] and the other cohort enrolled both patients with clinical stage T1-2 (71 percent) and T3 (21 percent) but all patients had normal bone scan findings.[140]

Of the seven cohorts in the PSA screening era, three cohorts (4 publications) enrolled patients with clinical stage T2 or less,[137,145,146,150] one cohort enrolled patients with "low-stage, low-grade disease,"[149] one cohort enrolled patients with any T stage N0/X, M0/X,[112] and the other two cohorts did not report or did not use clinical stage as part of patient eligibility criteria. The commonly used patient eligibility criteria were PSA (5 cohorts), age (4 cohorts), Gleason score (4 cohorts), and normal bone scan findings (4 cohorts). More details of each eligibility criterion in these seven cohorts are described in the following sections.

Age. Four cohorts (5 publications) reported age as part of patient eligibility criteria.[144-147,150] The age criterion was less than 75 years in the WW arm in SPCG-4 (Scandinavian Prostate Cancer Group Study Number 4)[144] and PIVOT (Prostate Cancer Intervention Versus Observation Trial),[146] and less than 85 years in the Watchful Waiting Study.[147] One cohort at Erasmus University hospital enrolled patients from the European Randomized Study of Screening for Prostate Cancer (ERSPC).[145,150] The ERSPC-screening protocol was carried over to become part of the patient selection criteria for this cohort, which required patients' age to be between 50 and 75 years old.[151]

Gleason score. Four cohorts (5 publications) used the Gleason score as part of patient eligibility criteria.[112,144,145,147,150] The Gleason score criterion was less than 8 in three cohorts.[112,145,147,150] The SPCG-4 cohort reported that "patients whose condition was diagnosed with an extended biopsy protocol were accepted if less than 25 percent of the tumor was Gleason grade 4 and less than 5 percent grade 5."[144]

Number of cores positive for cancer. None of the 13 unique cohorts used number of cores positive for cancer as part of patient eligibility criteria.

Percentage cancer involvement in each core. None of the 13 unique cohorts used percent cancer involvement in each core as part of patient eligibility criteria.

Prostate specific antigen. Five cohorts (6 publications) used PSA threshold as part of patient eligibility criteria.[144-148,150] Two criteria were used: PSA less than 50 ng/mL (4 cohorts)[144,146-148] and less than or equal to 15 ng/mL (1 cohort).[145,150]

Imaging. Four cohorts (5 publications) required normal bone scan findings as part of patient eligibility criteria.[144-146,148,150] One cohort also required patients to have normal chest radiograph findings to be eligible for the observational management program.[145,150]

Behavioral indicators. No behavioral indicator was used explicitly as a criterion for patient enrollment in the 13 unique cohorts.

Followup Protocols of Observational Management Strategies with Palliative Intent (Table 8)

Five of the six cohorts in the pre-PSA screening era included regular prostate acid phosphatase (PAP) testing and bone scan in the followup protocol.[138-142] The sixth cohort reported regular PSA and DRE in the followup protocol for patients who received no treatment after the introduction of PSA in 1990.[143] No information regarding the followup protocol in the pre-PSA screening era was provided. The monitoring frequency varied across these six cohorts (Figure 6).

All seven cohorts (8 publications) in the PSA screening era included regular PSA testing.[112,144-150] The monitoring frequency varied across the cohorts (Figure 6). Compared with AS cohorts (see previous section), rebiopsy was not commonly included in the followup protocol among WW cohorts.

Gleason score. None of the 13 unique cohorts used the Gleason score as part of followup protocols for patients who did not receive initial treatments.

Number of cores positive for cancer. None of the 13 unique cohorts used number of cores positive for cancer as part of followup protocols for patients who did not receive initial treatments.

Percentage cancer involvement in each core. None of the 13 unique cohorts used number of cores positive for cancer as part of followup protocols for patients who did not receive initial treatments.

PSA. Three cohorts formed in the pre-PSA screening era reported that PSA testing became part of followup protocol after PSA became available.[138,143,147] All six cohorts in the PSA screening era included regular PSA testing as part of followup protocol. However, rising PSA concentration alone was not used as a trigger for treatment in five cohorts.[112,144,146,147,149] The sixth cohort reported that "hormonal manipulation was demanded by the protocol when the PSA rose to 50 ng/mL."[148]

Imaging. Five cohorts in the pre-PSA screening era included regular bone scan in the followup protocol.[138-142] The monitoring frequency varied across the cohorts (Figure 6). One cohort also included regular chest and skeletal radiographs in the followup protocol.[142] Another cohort reported that computed tomography of the pelvis was conducted infrequently.[138]
Three cohorts (4 publications) in the PSA screening era included regular bone scans and chest radiographs in the followup protocol.[144-146,150] The monitoring frequency varied across the cohorts (Figure 6). Another cohort reported that all patients underwent "multiple bone scans" during followup.[148]

Behavioral indicators. No behavioral indicator was used explicitly in any of the followup protocols.

Table 7. Eligibility criteria for enrollment in protocols with palliative intent in chronological order of starting enrollment year

Center, Country [Pubmed ID] Enrollment years	Term used	Age (yr)	PSA (ng/mL)	Gleason score	# biopsy cores /% cores	Imaging	Stage	Behavioral indication (other than patients' choice or preference)
Howard University College of Medicine, US[138] [1600492] 1967-1989	EM /WW	–	–	–	–	–	Stage A and B	–
Orebro Medical Center, Sweden[139] [7933233] 1977-1984	Deferred treatment	Any (all patients >75 were not given any initial treatment from 1978-79)	–	–	–	–	T0-T2	–
Northern Stockholm region, Sweden[140] [17467883] 1978-1982	WW	–	–	–	–	Normal bone scan	71% T1-2 and 29% with T3	–
Freeman Hospital, UK[141] [3191340] 1978-1985	Deferred treatment	–	–	–	–	–	–	Patients without symptoms after initial outflow tract surgery or biopsy
Western General Hospital, UK[142] [8343901] 1978-1990	Deferred treatment	–	–	–	–	–	incidental (T0/stage A) or localized *T1/stage B1/B2)	–

Table 7. Eligibility criteria for enrollment in protocols with palliative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Term used	Age (yr)	PSA (ng/mL)	Gleason score	# biopsy cores /% cores	Imaging	Stage	Behavioral indication (other than patients' choice or preference)
Taichung Veterans Hospital, Taiwan[143] [12854876] 1983-1996	No treatment	–	–	–	–	–	T1a	–
SPCG-4, Finland, Sweden, and Iceland[144] [12226148] 1989-1999	WW	<75	<50	If diagnosed with an extended biopsy protocol, <25% of the tumor Gleason grade 4 and <5% grade 5	–	Bone scan negative	T0d, T1 or T2; T1c (after 1994)	–
Erasmus Univ. Hospital, Netherlands[145] [7544841] ⩽1990; 1993-2006[150]	No treatment /AS[150]	50-75[a151]	⩽15[150]	<8[150]	–	Bone scan and chest x-ray negative	T1c or T2[150]	No treatment decision was made by the urologist in discussion with the patient and his family, with respect to patient age, general health, clinical stage and patient preference.[145] All patients had estimated survival >1 yr.
Royal Marsden Hospital, UK[112] [15839912] 1993-2002	WW	–	Any	⩽7	–	–	any T stage, N0/X, M0/X	Unsuitable for RP typically because advanced age or comorbidities.
PIVOT, US [146] [18783735] 1994-2002	WW	⩽75	⩽50	Any	–	Bone scan negative	T1-T2/Nx/M0	–
Watchful Waiting Study, US[147] [14501381] 1998-2003	WW	<85	<50	<8	–	–	–	> 3-yr life expectancy, no history of any type of malignancy within the past 5 yr with the exception of non-melanoma skin cancer

Table 7. Eligibility criteria for enrollment in protocols with palliative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Term used	Age (yr)	PSA (ng/mL)	Gleason score	# biopsy cores /% cores	Imaging	Stage	Behavioral indication (other than patients' choice or preference)
Hospitals in Manchester region, UK[148] [11711356] NR	WW	–	<50	–	–	Bone scan negative	–	–
Univ. of Florida, US[149] [18263992] 2003-2006	EM	–	–	–	–	–	"Low-stage, low-grade disease"	Severe medical condition with a life expectancy of <10 yr

WW = watchful waiting; EM = expectant management; NR = not reported; DT = doubling time; mo = month(s); PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = Transrectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; US = ultrasound; yr = year(s); ED = erectile dysfunction; SPCG-4 = Scandinavian Prostate Cancer Group Study Number 4; PIVOT = Prostate Cancer Intervention Versus Observation Trial; VA = Veterans Affair

[a] ERSPC-screening protocol, carried over to AS selection criteria because patients were identified from ERSPC

[b] University Hospital of South Manchester, Withington Hospital, Christie Hospital; Hope Hospital

Table 8. Monitoring criteria in protocols of observational management strategies with palliative intent in chronological order of starting enrollment year

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Howard University College of Medicine, US[138] [1600492] 1967-1989	Every 3 mo for the first 5 yr, every 4-6 mo thereafter. Assessment included DRE, PAP and since 1985 a PSA was done.	√ (after 1985)	–	–	Annual bone scan; CT of the pelvis was used infrequently	–	PAP	Signs and/or symptoms of disease activity.
Orebro Medical Center, Sweden[139] [7933233] 1977-1984	Every 6-12 mo clinical exam, PAP, and bone scans. PSA only performed in the last few yr	√ (in the last few yr)	–	–	Bone scan	–	PAP	Patients were treated hormonally if disease progressed for they had symptoms of progression.
Northern Stockholm, Sweden[140] [17467883] 1978-1982	Every 3 to 6 mo for the first 2 yr and every 6 to 12 mo thereafter with DRE and PAP; annual rebiopsies during the first 4 yr	–	–	–	Bone scan every 12 to 18 mo	–	PAP	Treatment was offered if clinical progression with symptoms
Freeman hospital, UK[141] [3191340] 1978-1985	NR (“Disease progression was monitored”)	–	–	–	6-monthly bone scans (after 1983)	–	Acid and alkaline phosphatase	No treatment until symptomatic progression.
Western General Hospital, UK[142] [8343901] 1978-1990	Every 3 mo	–	–	–	Chest X-rays, skeletal X-rays and bone scans every 6 mo	–	√	Progression of disease (i.e., development of metastases (M1) or elevation of PAP to > 2 u/l) and/or development of symptoms.

Table 8. Monitoring criteria in protocols of observational management strategies with palliative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Taichung Veterans hospital, Taiwan[143] [12854876] 1983-1996	Every 3-6 mo PSA and DRE (after 1990, introduction of PSA)	√ (after 1990)	–	–	–	–	–	No treatment until there was evidence of cancer progression.
SPCG-4, Finland, Sweden, and Iceland[144] [12226148] 1989-1999	Every 6 mo in the first 2 yr, then every 1 yr	√	–	Rebiopsy was not routinely undertaken[152]	A bone scan and chest radiograph were obtained annually until 1997; thereafter, chest radiographs were obtained annually for the first 2 yr	–	Hemoglobin, creatinine, alkaline phosphatase	Adjuvant local or systemic treatment was not given. TURP was as a treatment for local progression.[a]
Erasmus Univ. hospital, Netherlands[145] [7544841] ≤1990; 1993-2006[150]	Usually Followed clinically every 6 mo; Follow-up regimens varied among local practices[150]	√	–	–	"Bone scan and chest x-ray were repeated regularly"	–	Alkaline phosphatase	Local[b] and metastatic progression were evaluated. Subjective progression, like obstructive micturition or pain, was considered for treatment decisions.[145] The authors reported that of 13 patients with progression, 6 started treatment (5 for subjective symptoms; 1 for objective progression only). The authors also reported that PSA progression may serve as a trigger point to treatment.[150]

Table 8. Monitoring criteria in protocols of observational management strategies with palliative intent in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Royal Marsden Hospital, UK[112] [15839912] 1993-2002	Every 6 mo PSA and DRE	√	–	–	–	–	–	Symptomatic prostate cancer progression
PIVOT, US [146] [18783735] 1994-2002	Every 6 mo PSA	√	–	–	Bone scan every 5 yr	–	–	Discouraged treatment for asymptomatic progression (eg, per PSA)
Watchful Waiting Study, US[147] [14501381] 1998-2003	Every 3 mo PSA	√	–	–	–	–	–	Developing progressive disease
Hospitals in Manchester, UK[148] [11711356] NR	Every 6 mo PSA	>50	–	–	All patients underwent "multiple bone scans" (all negative),	–	–	Hormonal manipulation was demanded by the protocol when the PSA rose to 50 ng/mL.
Univ. of Florida, US[149] [18263992] 2003-2006	Every 3 mo PSA and DRE; Repeat biopsy about 6 mo after the initial diagnosis.	√	–	–	–	–	–	Cancer progresses or symptoms become imminent.

WW = watchful waiting; EM = expectant management; NR = not reported; DT = doubling time; mo = month(s); PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = Transrectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; US = ultrasound; yr = year(s); ED = erectile dysfunction; SPCG-4 = Scandinavian Prostate Cancer Group Study Number 4; PIVOT = Prostate Cancer Intervention Versus Observation Trial; VA = Veterans Affairs

√ = item was used as part of monitoring strategy but explicit criteria were not defined

– = item was not used or not reported as part of monitoring strategy

[a] Local progression was defined as a palpable transcapsular tumor growth; symptoms of obstruction of the flow of urine that necessitated intervention, or both.

[b] Local progression was defined as symptoms (subjective), increase in T category, increase in prostate size on DRE by 25%, or increase in ultrasound measured volume >40%.

Figure 6. Followup frequencies of 13 unique cohorts of observational management strategies with palliative intent[a]

	Year	1				2				3				4				5			
WW cohort	**Month**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27**	**30**	**33**	**36**	**39**	**42**	**45**	**48**	**51**	**54**	**57**	**60**
Pre-PSA screening era																					
HUCM[138]	PAP; PSA (after 1985)	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*
	DRE	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*
	Bone scan				√				√				√				√				√
Orebro Medical Center [147]	PAP			√				√				√				√				√	
	DRE			√				√				√				√				√	
	Bone scan			√				√				√				√				√	
Northern Stockholm[140]	PAP		√		√		√		√			√				√				√	
	DRE		√		√		√		√			√				√				√	
	Rebiopsy				√				√				√				√				√
	Bone scan					√				√				√							
Freeman Hospital[141]	PAP	monitoring frequency was not reported																			
	DRE																				
	Bone scan		√		√		√		√		√		√		√		√		√		√
Western General Hospital[142]	PAP	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE																				
	Bone scan; X-ray		√		√		√		√		√		√		√		√		√		√
Taichung Veterans Hospital[143]	PSA (after 1990)		√		√		√		√			√				√				√	
	DRE		√		√		√		√			√				√				√	
	Bone scan																				
PSA screening era																					
SPCG-4[137]	PAP; PSA		√		√		√		√				√				√				√
	DRE																				
	Rebiopsy	not routine																			
	Bone scan; X-ray				√				√				√				√				√
Erasmus University Hospital[145]	PSA		√*		√*		√*		√*		√*		√*		√*		√*		√*		√*
	DRE																				
	Bone scan; X-ray	"repeat regularly"																			
Royal Marsden Hospital[112]	PSA		√		√		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√		√		√
	Bone scan																				
PIVOT[146]	PSA		√		√		√		√		√		√		√		√		√		√
	DRE																				
	Bone scan																				√
Watchful Waiting Study[147]	PSA	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE																				
	Bone scan																				

Figure 6. Followup frequencies of 13 unique cohorts of observational management strategies with palliative intent[a] (continued)

	Year	1				2				3				4				5			
WW cohort	Month	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60
Hospitals in Manchester region[148]	PSA		√		√		√		√		√		√		√		√		√		√
	DRE																				
	Bone scan	Probably not routine [b]																			
University of Florida[149]	PSA	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Rebiopsy		√																		

Legends: √* = every 3 mo for the first 5 yr; every 4-6 mo thereafter; √** = usually every 6 mo but followup regimens varied among local practices. Merged cell represents a range of followup frequency; for example, a merged cell of 3 and 6 mo with a check mark in the middle of the merged cell means a followup frequency of 3 to 6 mo. PAP = prostate acid phosphatase HUCM = Howard University College of Medicine; Northern Stockholm = Northern Stockholm region in UK; SPCG-4 = Scandinavian Prostate Cancer Group Study Number 4 trial; PIVOT= Prostate Cancer Intervention Versus Observation Trial; Hospitals in Manchester region= University Hospital of South Manchester, Withington Hospital, Christie Hospital; Hope Hospital

[a] Cohorts of watchful waiting in chronological order of starting enrollment year. See Table 8 for more detailed monitoring criteria in each cohort.

[b] The authors reported that all patients underwent "multiple bone scans" during followup and all had normal findings.

Observational Management Strategies with Unclear Treatment Intent

Six cohorts reported followup protocols but did not report triggers for treatment of prostate cancer, so it is unclear what observational management strategies were used to indicate the need for treatments in those patients who did not receive initial treatments (Table 9). Various terms were used to describe the observational management strategies in these cohorts, including "no treatment," "expectant management," "watchful waiting," and "active surveillance." The eligibility criteria and followup protocol of these cohorts were summarized in Tables 10 and 11. None of these cohorts used parameters that have not been previously described.

Table 9. Cohorts that did not report triggers for treatment of prostate cancer

Country	Center or Study Name
US	Kansas City Veterans Affairs Hospital University of North Carolina
Canada	British Columbia Cancer Agency Princess Margaret Hospital
Japan	Kagawa Medical University (1990-1998) Kitasato University Hospital

Summary

We identified 16 unique cohorts reporting formal protocols to monitor triggers for curative treatment of prostate cancer. The eligibility criteria for patient selection and followup protocols were heterogeneous.

Among these cohorts, the most commonly used parameter as part of patient eligibility criteria was Gleason score (12 cohorts), no higher than Gleason 6 or 7. More recently, Gleason patterns were also used in some of these AS cohorts, such as no higher than Gleason 3+3 or 3+4. All 16 cohorts enrolled only patients with clinical stages T2 or less and included regular PSA testing in the followup protocol. PSA and Gleason score were the most commonly used parameters as part of monitoring criteria for disease progression. Generally, progression in Gleason was defined as a Gleason score or pattern greater than those used in the eligibility criteria for AS. Regularly scheduled rebiopsy was also a common parameter in the AS followup protocol. Large variation exists in terms of the definitions of disease progression, and the frequencies of AS monitoring protocols.

In contrast to the above AS cohorts, less variability exists in terms of the definitions of eligibility criteria for patient selection and followup protocols among the 13 cohorts of other observational management strategies. All such cohorts used only symptomatic progression as triggers for treatment; thus we labeled these observational management strategies as having primarily palliative treatment intent. Regular bone scan schedule was commonly included in these followup protocols. Rebiopsy was typically not used in these strategies; imaging tests were more commonly used to track disease progression.

Implicit in the Key Question is a comparison between AS and other observational strategies in the modern PSA era. Thus, we compared the 16 unique cohorts reporting formal protocols to monitor triggers for curative treatment with the seven unique cohorts of other observational strategies with primarily palliative intent in the PSA screening era. Enrollment into AS protocols more commonly used Gleason score as a threshold than other observational strategies. They also used the number and percentage of cores positive for cancer as a threshold while none of the

other strategies used these factors. Both sets of strategies generally used some sort of PSA criteria, but the thresholds in AS were generally lower (10-15 ng/mL) than the other observational strategies (15 or 50 ng/mL). AS protocols had more clearly defined followup than other observational strategies, with explicit indications for curative treatment including increase in Gleason scores, number and percentage of positive cores (on rebiopsy), and/or PSA velocity. AS protocols generally did not include imaging in their followup protocols. In contrast, other observational strategies typically included imaging in their followup, specifically bone scan and chest radiography. They also generally did not use rebiopsy but they did use PSA in their followup. Comparison of the followup frequencies between AS and other observational strategies (Tables 2.3 versus Table 2.10) showed that PSA testing and DRE were common in both strategies, but somewhat more frequent with AS protocols, at least within the first year of followup.

Table 10. Protocols that did not report information on triggers for intervention in chronological order of starting enrollment year

Center, Country [Pubmed ID] Enrollment years	Term used	Age (yr)	PSA (ng/mL)	Gleason score	# biopsy cores /% cores	Imaging	Stage	Behavioral indication (other than patients' choice or preference)
Kagawa Medical Univ., Japan[153] [10765093] 1990-1998	EM;WW	–	"elevated PSA"	⩽6	1-2 positive cores per 6 sextant cores; ⩽50% involvement of any positive core	–	–	–
Kitasato Univ. Hospital, Japan[154] [11851612] 1991-2000	WW	–	–	–	6 sextant biopsy	–	"clinically localized prostate cancer"	–
Univ. of North Carolina, US[155] 1991-1996	EM	–	–	–	–	–	T1c	–
Princess Margaret Hospital, Canada[156] [21211899] 1995-2010	AS	–	<10	<6	⩽3 positive biopsy cores (<50% of a core involved at initial diagnostic biopsy); fist-time biopsies consisted of 6 cores before 2001 and 11 cores after 2001.	–	T1c-T2a	–
BCCA, Canada[157] [9445192] NR	WW	–	–	–	–	–	–	Patient wish (37%), reduced life expectancy due to medical problem (19%), physician recommendation (42%); relative contraindication to RT (2%)
Kansas City VA, US[158] [21172105] 2004-2009	AS[a]	–	<20	<6	<20% positive biopsy		⩽ T2	–

WW = watchful waiting; EM = expectant management; NR = not reported; DT = doubling time; mo = month(s); PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = transrectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; US = ultrasound; yr = year(s); BCCA = British Columbia Cancer Agency

[a] AS criteria were created explicitly for the analyses only, this is not a prospective AS cohort.

Table 11. Monitoring parameters in cohorts that did not report information on triggers for intervention in chronological order of starting enrollment year

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Kagawa Medical Univ., Japan[153] [10765093] 1990-1998	NR	PSA DT based on 1st PSA >1 mo after biopsy. ⩾3 values at intervals ⩾1 mo apart for >6 mo.	–	–	–	–	–	NR
Kitasato Univ. Hospital, Japan[154] [11851612] 1991-2000	"a DRE," generally seen every 3-6 mo "as clinical circumstances dictated."	–	–	–	Annual bone scan	–	–	NR
Univ. of North Carolina, US[155] 1991-1996	PSA at 3 mo; then every 6 mo	Biochemical progression: PSA level increase in 3 consecutive measurements and the total increase was > 5 ng/mL	–	–	–	–	Hematocrit and creatinine every 6 mo	NR
Princess Margaret hospital, Canada[156] [21211899] 1995-2010	For the most part, PSA every 3mo for 2 yr and every 6 mo in stable patients; DRE every 6 mo; a confirmatory biopsy within 12 mo and then every 2–3 yr until the patient reached 80 yr of age or refused treatment[a]	√	–	Repeat biopsies consisted of 10 cores before 2001 and 15-16 cores after 2001. >3 cores or any core involvement >50%	–	–	–	NR Note: Pathologic progression was devalued, defined as increased grade, increased number of cores to more than 3 or any core involvement >50%.

Table 11. Monitoring parameters in cohorts that did not report information on triggers for intervention in chronological order of starting enrollment year (continued)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions			
BCCA, Canada[157] [9445192] NR	PSA generally every 3-6 mo as needed			√	–	–		–	–	–	NR
Kansas City VA, US[158] [21172105] 2004-2009	PSA every 3 mo and a repeat TRUS guided prostate biopsy at 1 yr			√	–	All biopsies were performed using a standard 12-core biopsy scheme, but increased number if larger glands		–	–	–	NR

WW = watchful waiting; EM = expectant management; NR = not reported; DT = doubling time; mo = month(s); PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = Transrectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; US = ultrasound; yr = year(s); BCCA = British Columbia Cancer Agency

√ = item was used as part of monitoring strategy but explicit criteria were not defined

– = item was not used or not reported as part of monitoring strategy

[a] The authors reported that five physicians in a nonstandardized fashion followed patients, although a relatively similar pattern of care was provided.

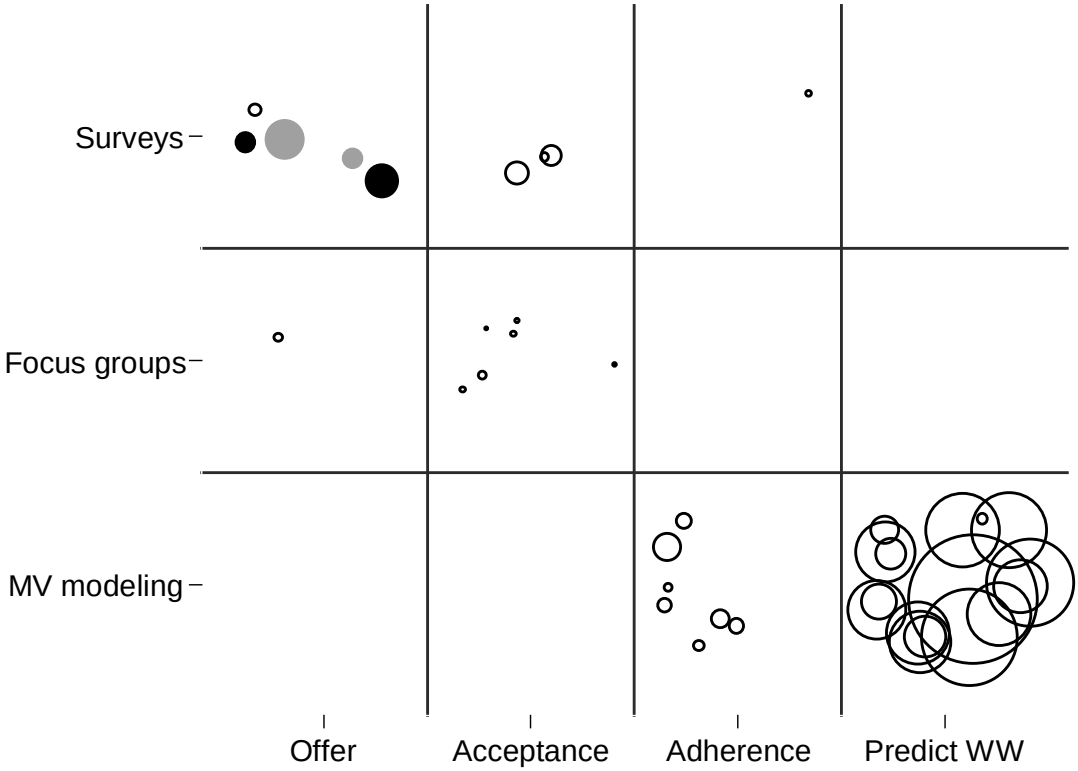
Key Question 3. What factors affect the offer of, acceptance of, and adherence to active surveillance?

For this Key Question, eligible studies included: (1) multivariable database analyses of predictors for the offer of, acceptance of, and adherence to AS (or WW), (2) survey or interview type studies addressing the same issues, and (3) experimental studies that examined a factor of interest addressing the same issues, when applicable (e.g., the effect of decision aids on the acceptance of AS). Eligible studies reporting multivariable analyses had to adjust for age and disease stage or risk. We excluded studies in which AS/WW was not analyzed separately from nonaggressive treatments like ADT. Similarly, for survey or interview type studies, only those including men with prostate cancer and reporting data directly relevant to AS/WW were reviewed. Of note, the outcomes of many of the studies were either treatment with an observational strategy or interruption (cessation) of the observational strategy. Studies generally did not directly analyze the offer of, acceptance of, and adherence to AS.

Figure 7 summarizes the studies reviewed for Key Question 3. Twenty-five studies reported multivariable analyses of the association between different physician or patient factors, delivery system, and the offer of, acceptance of, and adherence to, AS or WW.[66,77,159-181] These analyses were mainly conducted on the CaPSURE or SEER databases. In addition, 16 survey or interview type studies explored similar associations.[122,182-196] No experimental study specifically examined factor(s) addressing the offer of, acceptance of, and adherence to AS. However, one relevant systematic review detailed the use of decision-making tools and aids in the management of men with prostate cancer.[197] As described in the Methods chapter, the included studies were those initially identified in our search for publications concerning active surveillance, and from references in relevant reviews. We did not do a targeted search for the specific factors of interest.

It should be noted that among this group of studies, only two specifically targeted men who were put on an active monitoring protocol with triggers for curative treatments.[172,173] The remainders were analyses of men who were either not treated or not initially treated. We could not determine whether they were on an active monitoring protocol with triggers for curative treatments.

Figure 7. Evidence map for key question 3



Studies considered for Key Question 3 are depicted as circles with size proportional to the study sample size, stratified by design (focus groups, surveys, or studies reporting multivariable prediction models) and outcomes assessed ("offer" of, "acceptance" of, "adherence" to observational management strategies or prediction of having received observational management). One study that reported factors influencing the acceptance of randomization versus non-randomized choice of treatment is not shown. Hollow circles represent studies including prostate cancer patients only; black circles represent studies of physicians only; gray circles represent studies including both patients and physician. Placement within each box is random. MV = multivariable; WW = watchful waiting.

Only the qualitative results of the multivariable analyses are described in this section. Because the reviewed studies used heterogeneous coding schemes for their predictors (for example, age was used as a continuous variable in some studies but as a nominal [discrete] variable in others) and adjusted for varying sets of confounding variables, meaningful comparison of effect sizes across studies is precluded. For these reasons, we do not present effect estimates for each predictor of interest from these analyses (such as odds ratios for predicting treatment received, or hazard ratios for WW interruption), although detailed quantitative information is available in Appendix Tables C3.1 to C3.3.

It should also be noted that the common method for reporting "adherence to AS" in the literature is the "interruption of AS" to seek definitive treatments and we follow this convention in our review. A man could interrupt AS to seek curative treatments for several reasons, among which: 1) the person meets some criteria on AS protocol indicative of disease progression that would call for curative treatment, 2) the person does not meet criteria for curative treatment (i.e., continued surveillance is indicated), but due to personal preference, he decides to stop AS and pursue curative treatment, and 3) the person decides to forgo present or future curative treatment (e.g., because of advanced age or new comorbidities) and switches to WW. The first reason would commonly be considered "adherent" (the person is following the protocol) and the latter two "not adherent" (he chose to discontinue the AS protocol even though there was no indication

of disease progression). However, the studies reviewed largely ignored this distinction or had insufficient data to make this distinction and assessed adherence by reporting "interruption of AS" to seek definitive treatments.[a] The one exception was those studies that assessed patient anxiety leading to curative treatment despite not meeting disease progression criteria.

Physician Factors (Appendix Tables C3.1–3.3)

Primary Care

Offer of AS. No study specifically examined how the involvement of a primary care physician in the decisionmaking process might affect the offer of AS.

One survey of 381 New Zealand general practitioners given clinical vignettes reported that 45 percent of general practitioners would recommend WW if a patient's life expectancy was less than 10 yr, but only 3 percent would recommend WW if a patient's life expectancy was more than 10 yr.[184]

Acceptance of AS. No study specifically examined how a patient's primary care physician might affect the acceptance of AS.[b]

Three survey/interview-type studies (sample sizes were 25,[183] 102,[187] and 185[122]) reported that physician recommendation (urologists or radiation oncologists) was the most influential factor in a patient's decision (30 percent in Holmboe 2000;[187] 73 percent in Gorin 2011;[122] no quantitative data were available in Davison 2009[183]) to elect or not elect AS. Two other surveys of men with prostate cancer (sample sizes were 654[185] and 231[190]) reported that physician recommendation (urologists, radiation oncologists, and others) was most influential in reaching a treatment decision (51 percent[185] and 57 percent[190]). However, one multivariable analysis of men in the Prostate Cancer Research International: Active Surveillance study (PRIAS) reported that patients who perceived that physician played the most important role in shared decision-making process also had higher decisional conflict (more doubts) regarding the patient's choice of AS ("This suggests that men who perceive that they have actively participated in the treatment decision-making process have fewer doubts regarding their treatment decision").[181]

Adherence to AS. No study specifically examined how a patient's primary care physician might affect adherence to AS. However, one survey of 53 men on AS who ultimately received treatment reported that 81 percent believed that treatment was favored by their physicians (urologists, radiation oncologist, medical oncologist, primary care physician), which was the primary cause of the change in plan for AS.[186] In contrast, physician notes revealed that for only 24 percent of the patients was there documentation that the physician recommended treatment due to clinical or biochemical evidence of tumor progression, leading to the study's conclusion that physicians more often perceive that patients themselves initiated the treatment decisions.

---

[a] We do not know the proportion of non-adherence. We do know that the proportion of patients on AS who went on to receive active treatments ranged from 16 to 54 percent with a median followup of 1 to 7 years (see Appendix Table C5.1).

[b] We are aware of one age- and comorbidity-stratified analysis of 85,088 men with clinically localized disease identified from the SEER-Medicare database which concluded that men who saw a primary care physician after diagnosis were more likely to have AS/WW than those who did not (e.g., for men aged 75-79 years with comorbidity index of $\geqslant$2, 76.5 vs. 12.2 percent, see original paper for results from other categories).[198] This study did not meet our multivariable analysis inclusion criterion.

Diagnosing Physician

Offer of AS. No study specifically examined how the involvement of the diagnosing physician in the decisionmaking process might affect the offer of AS. However, one survey of 185 men already on AS reported that AS was offered by 36 percent of the physicians who had made the initial diagnosis.[122]

Acceptance of AS. No study or survey specifically addressed how the involvement of the diagnosing physician might affect the acceptance of AS.

Adherence to AS. No study specifically examined how the involvement of the diagnosing physician might affect the adherence to AS.

Consultant—2nd Opinion

Offer of AS. No study specifically examined how the involvement of a consulting physician for a second opinion in the decisionmaking process might affect the offer of AS. However, a description of interviews with 18 couples in which the men were recently diagnosed with early stage prostate cancer and had not yet decided on a treatment reported that "the urologist had recommended seeking a second opinion and indeed had offered to facilitate such a referral [for several couples]. None followed through with this suggestion...The fact that the urologist [had recommended seeking] a second opinion only further reinforced their trust and confidence...."[194]

In an interview of 108 men in Australia with newly diagnosed localized prostate cancer, concerning their urological consultation, 71 percent reported that their urologists discussed WW (versus 92 percent for RP and 87 percent for RT).[192]

One survey of 200 urologists querying their preferences for treatments for men with localized prostate cancer and few comorbidities reported that 67 percent preferred RP, 29 percent preferred RT, and 4 percent preferred WW.[193] The same study also surveyed 780 men with all stages of prostate cancer and reported divergent opinions (patient versus physician) on whether treatment options were discussed: 20 percent of the men versus 1 percent of the urologists felt that treatment options were not discussed. It should be noted, however, that the urologists in the survey were not necessarily the surveyed patients' own urologists.

One survey of 238 men with newly diagnosed prostate cancer and their 25 urologists reported on their office encounters.[182] Ninety-five men presented for an initial consultation, and 143 men presented for a second opinion visit. The urologists recommended 0.52 more treatment options (SE 0.19, $P < 0.001$) in the initial consultation setting than in the second opinion visit setting. For men with low-risk disease, 25 percent of the urologists recommended AS and 77 percent recommended RP in the initial consultation setting, but only 16 percent recommended AS and 91 percent recommended RP in the second opinion visit setting. The survey also reported a discrepancy between what the physicians recommended and what the patients heard: in those patients for whom the urologists recommended RP, 67 percent reported receiving the recommendation; in those patients for whom the urologists recommended RT or ADT, only about 25 percent of the patients reported receiving the recommendation.

Acceptance of AS. No study or survey specifically addressed how the involvement of a consulting physician might affect the acceptance of AS.[c]

[c] We are aware of one age-stratified analysis of 85,088 men with clinically localized disease identified from the SEER-Medicare database which concluded that for all age groups, men who saw urologists either with or without

Adherence to AS. No study or survey specifically addressed how the involvement of a consulting physician might affect adherence to AS.

Clinical Factors

Offer of AS. One survey of 1063 urologists and radiation oncologists reported that about 10 to 20 percent would recommend WW for a patient with a PSA of around 5 ng/mL and a Gleason score of around 4 or 5 (the given scenario was a 65 years old man in good health, with negative DRE and no evidence of nonlocalized disease), but almost none would recommend WW for those with higher PSA or Gleason scores (Fowler 2000).[188] The responses of urologists and radiation oncologists did not differ significantly.

Acceptance of AS

Age. Twelve multivariable analyses provided results for age with respect to AS/WW.[66,77,159,161,162,166,171,174,175,177,179,180] All reported men who were older (generally aged 65 to 75 years) had an increased probability of receiving AS/WW versus active treatments.

Comorbidities. Five multivariable analyses reported that men with an increased number of comorbidities also had an increased probability of receiving AS/WW versus active treatments.[66,77,159,164-166] Two multivariable analyses did not find an association.[177,180]

Gleason score. Three multivariable analyses reported that men with a higher Gleason score also had a decreased probability of receiving AS/WW versus any other treatment.[77,180] or RP.[171]

Histopathology. Two multivariable analyses provided results for histopathology with respect to AS/WW.[159,166] Both reported that men with well-differentiated, as compared to either moderately or poorly differentiated, prostate tumors had an increased probability of receiving AS/WW versus other treatments.

Stage. Three multivariable analyses reported that men with higher stage disease (local versus in situ;[199] T2 versus T1[171,177]) had a decreased probability of receiving AS/WW versus active treatments.

PSA. Three multivariable analyses reported that men with increased PSA had a decreased probability of receiving AS/WW versus active treatment.[77,171,177] One multivariable analysis of men with high risk but nonmetastatic prostate cancer reported that higher PSA (>20 vs. ≤10 ng/mL) was associated with increased probability of receiving AS/WW versus RP.[179]

Risk groups. Two multivariable analyses reported that men assessed as D'Amico low-risk[200] (versus intermediate- or high-risk) had an increased probability of receiving AS/WW versus active treatments.[66,161]

---

medical oncologists were more likely than those seen by both urologists and radiation oncologists to receive AS/WW (all ages combined: 26.3 percent or 34.1 percent vs. 4.6 percent, respectively; see original paper for age-stratified results).[198] This study did not meet our multivariable analysis inclusion criterion.

Adherence to AS

Age. Four multivariable analyses analyzed age with respect to interruption of AS/WW.[164,168,175,176] Two found that men who were younger had an increased probability of receiving definitive treatments,[168,175] and two found that age was not a factor.[164,176]

Comorbidities. Three multivariable analyses reported that the number of comorbidities was not associated with interruption of AS/WW.[164,175,176]
Gleason score. Five multivariable analyses reported that Gleason score at diagnosis was not associated with interruption of AS/WW.[168,172,174-176]

Histopathology. No study analyzed histopathology to describe tumor differentiation with respect to interruption of AS.

Stage. Three multivariable analyses reported that men with higher stage disease (T2 versus T1) had an increased probability of interruption of AS/WW to seek secondary treatment.[168,175,176] One analysis reported that disease stage (T2 versus T1) was not a significant predictor of the receipt of curative treatment.[174]

PSA. Two multivariable analyses reported that men with increased PSA at diagnosis also had an increased probability of interruption of AS/WW to seek active treatments.[168,175] One analysis reported that neither initial PSA nor PSA density was predictive of men who interrupted AS to seek radical treatment, but the free to total PSA ratio was predictive of the probability of interruption of AS.[172] One analysis reported that initial PSA was not predictive of men who interrupted AS to seek definitive treatment, but that a short PSA doubling time was (< 2 yr vs. 2 to 5 yr).[176] Two other multivariable analyses also reported that increased PSA velocity during followup was predictive of interruption of AS to seek active treatments.[164,174]

Risk groups. One multivariable analysis reported that patients assessed as D'Amico low-risk (versus intermediate- or high-risk) had a decreased probability of interruption of AS/WW to seek active treatment,[167] while two reported that risk classification was not a significant predictor in the interruption of AS/WW.[164,168]

Patient Factors

Family Involvement

Offer of AS. No study or survey specifically addressed how family involvement might affect the offer of AS.

Acceptance of AS. One survey reported that 19 percent of 654 men,[185] and another survey 9 percent of 231 men,[190] mentioned that advice from family and friends was the most influential factor in reaching a treatment decision. In a content analysis of focus group or interview discussion including a total of 44 men with localized prostate cancer, 20 men reported relying on influential others (an individual whose illness experience and/or story had explicit influence on the participant's treatment decision) to make a treatment decision.[191] Of these 20 men, this influential other caused one man to consider WW more strongly and one to more likely reject

WW. One open-ended interview of 102 men with localized disease reported that one of the reasons for not electing WW was that their families were against that option (4 percent). Other reasons cited were fear of consequences for not selecting WW (64 percent) and perceived elevated risk because of increased PSA or Gleason score (12 percent).[187]

Adherence to AS. No study or survey specifically addressed how family involvement might affect adherence to AS.

Personal Preferences

Offer of AS. No study or survey specifically addressed how personal preferences might affect the offer of AS.

Acceptance of AS. One analysis of the ProtecT trial compared 180 men who refused randomization, but instead selected AS (i.e., men who did not participate in the trial) with 138 men in the trial who were randomized to AS.[173] The analysis found that men with increased baseline anxiety (per unit increase on the Hospital Anxiety and Depression scale; adjusted OR 0.93; 95 percent CI 0.87, 0.99; P = 0.04) and lower SES (per decrease in SES from I [high] to V [low]; adjusted OR 0.68; 95 percent CI 0.49, 0.96; P = 0.03) had decreased probability of selecting AS and refusing randomization (i.e., these men did not proactively seek AS but preferred randomization for AS vs. active treatment).

Four survey/interview type studies (sample size 25,[183] 50,[189] 185,[122] and 768[195]) reported that concern for treatment side effects (impotence [44 percent] and incontinence [48 percent];[122] side effects (11 percent, type not specified);[195] no quantitative data were available in two studies[183,189]) was one reason that patients elected AS/WW.

Some other reasons cited in an interview of 21 men with localized disease two of whom chose WW were to maintain current quality of life and no need for active therapy because cancer was small and slow-growing.[196] The authors further observed that "for most men [including both who chose WW and those who did not], both black and white, treatment decision making occurred within an emotional context of fear and uncertainty and without systematic use of information."

One multivariable analysis reported that the desire to avoid side effects or having current bowel problems were predictive of the choice of WW versus other treatments or undecided.[178] One multivariable analysis reported that having urinary dysfunction was predictive of choosing WW over RP, while having sexual dysfunction was predictive of choosing RT over WW.[177] Another multivariable analysis reported that having other urinary conditions (besides the primary urinary dysfunction) was also predictive of choosing WW over RP.[165]

Adherence to AS. One multivariable analysis reported that increased anxiety in men was associated with an increased probability of interruption of AS.[164]

Risk Perceptions

Offer of AS. No study or survey specifically addressed how risk perceptions might affect the offer of AS.

Acceptance of AS. One qualitative description of interviews conducted in 25 men with low-risk prostate cancer reported that physician description of prostate cancer affects patient perception of

the seriousness of the condition as well as treatment choice.[183] One survey of 654 men with early stage prostate cancer reported that men who chose RP over RT or WW perceived prostate cancer as a significantly more serious disease.[185] Another survey of 102 men with localized prostate cancer reported that fear of consequences was the most common reason for not selecting WW.[187]

Adherence to AS. No study or survey specifically addressed how risk perceptions might affect the adherence to AS.

Family History

Offer of AS. No study or survey specifically addressed how family history might affect the offer of AS.

Acceptance of AS. No study or survey specifically addressed how family history might affect the acceptance of AS.

Adherence to AS. Two multivariable analyses reported that family history was not a significant factor in predicting interruption of AS/WW.[175,176]

Social Support

Offer of AS. No study or survey specifically addressed how social support might affect the offer of AS.

Acceptance of AS. Four multivariable analyses reported that not being married or in a permanent relationship was associated with an increased probability of receiving WW versus active treatments.[77,160,166,180] One analysis found that marital status was not a factor in predicting receiving WW.[177]

A report of interviews with 18 couples in which the men were recently diagnosed with early stage prostate cancer and had not yet decided on a treatment demonstrated the complexity of reaching a treatment decision.[194] The authors concluded that couples ruled out options based on both formal (provided by the physicians) and informal (provided by family and friends) information, and that they also "considered both their own individual histories and concerns and their shared life experiences." Of the 18 couples referred by urologists, only one couple elected watchful waiting. The authors further stated that "'Doing nothing' was ultimately rejected for the certainty [the couples] perceived to be associated with it: certain death, feared to be slow and painful."

Acceptance of AS. One multivariable analysis reported that marital status was not associated with time to interruption of AS.[164]

Other Factors That Could Affect the Offer of, Acceptance of, or Adherence to AS

Income or socioeconomic status. Three multivariable analyses examined income level with respect to AS/WW.[66,177,199] One reported that less than $30,000 annual income (versus at least $40,000) in men with prostate cancer was associated with an increased probability of receiving AS/WW versus other treatments.[199] The remaining two reported that income was not a significant factor in predicting the choice of AS/WW versus other treatments.[66,177]

One multivariable analysis reported that men in higher socioeconomic strata (versus lower strata) had a decreased probability of receiving AS/WW versus active treatments.[77,159]

Education. Five multivariable analyses examined education level with respect to AS/WW.[66,161,166,167,177] Three reported that education was not a significant factor in predicting men receiving AS/WW.[66,161,177] One reported that men who resided in census tracts with fewer residents who had a high school education (versus more) had an increased probability of receiving AS/WW versus other treatments.[166] One reported that college graduates (versus non-college graduates) had an increased probability of interrupting AS/WW to seek active treatments.[167]

Race. Ten multivariable analyses examined race/ethnicity with respect to AS/WW.[77,159,164,166,168-170,176,177,180] Six found that race/ethnicity was not a significant factor in selecting AS/WW or in the decision to interrupt AS/WW and seek definitive treatments.[77,164,168-170,176] Four analyses reported that blacks were more likely than whites to receive AS/WW versus active treatments.[159,166,177,180]

One survey of 231 men with prostate cancer in North Carolina reported that there was no significant difference between blacks and whites as to whether the option of WW was discussed with their physicians (48.7 percent versus 56.1 percent).[190]

Delivery System

Economic Incentives and Disincentives

Offer of AS. No study or survey specifically addressed how economic incentives and disincentives might affect the offer of AS.

Acceptance of AS

Insurance Type (HMO, Military, Private). Two multivariable analyses reported that having Medicare (versus private insurance,[66] or private or Veterans Administration (VA) insurance[160]) increased the probability of receiving AS/WW (versus active treatments). One analysis reported that having Medicare plus supplement (versus private) also increased the probability of AS/WW (versus RP).[179] One analysis reported that having preferred provider organization (PPO) or health maintenance organization (HMO) coverage decreased, and that having VA insurance increased, the probability of receiving AS/WW versus RP.[163] It also reported that Medicare supplemented with fee-for-service, HMO, or PPO coverage decreased the probability of receiving AS/WW versus RP.

Adherence to AS. One multivariable analysis reported that insurance status was not a significant factor in predicting interruption of AS/WW.[164]

Availability of Technology

No study or survey specifically addressed how the availability of technology might affect the offer or acceptance of, or adherence to, AS.

Geographic Location

Offer of AS

Small area variation. No study or survey specifically addressed how small area variation might affect the offer of AS.

Regional variation. No study or survey specifically addressed how regional variation might affect the offer of AS.

Urban vs. rural. One survey of 231 men with prostate cancer reported that there was no significant difference between urban and rural residents in North Carolina as to whether the option of WW was discussed with their physicians (51.9% vs. 53.7%).[190]

Acceptance of AS

Small area variation. No study or survey that specifically addressed how small area variation might affect the acceptance of AS. One study, however, did report that there was a wide variation in the selection of AS/WW across 36 practice sites in the United States (ranging from 0 to 28 percent) and that this variation was not explained by known patient factors.[69] The variation remained after restricting the analysis to men with low-risk disease. In a multivariable analysis, the proportion of variation for AS/WW among men with low-risk disease attributable to practice site was 21 percent (95 percent CI 0.11, 0.37).

Regional variation. One multivariable analysis claimed that men who resided in New Jersey versus those in California (excluding San Francisco-Oakland, San Jose, and Los Angeles) had an increased probability of receiving AS/WW versus any other treatments.[77] No significant differences were found between men in California (excluding San Francisco-Oakland, San Jose, and Los Angeles) and men in other registries (San Francisco-Oakland, San Jose, Los Angeles, Seattle, Detroit, Atlanta, Iowa, New Mexico, Utah, Louisiana, Connecticut, and Alaska natives). However, another multivariable analysis reported that men in Northeast had a decreased probability of selecting AS/WW (versus active treatments) compared with men in greater California (excluding San Francisco, San Jose-Monterey, and Los Angeles).[180]

Urban vs. rural. One multivariable analysis reported that men who resided in urban areas (versus rural areas) had a decreased probability of receiving AS/WW versus RP or RT.[159] One survey of 231 men with prostate cancer in North Carolina reported that there was a difference in whether physician recommendation was the most influential factor in the treatment decision between urban and rural residents (62.3 percent versus 43.9 percent, respectively; P=0.004). [190]

Adherence to AS

Small area variation. No study or survey specifically addressed how small area variation might affect adherence to AS.

Regional variation. No study or survey specifically addressed how regional variation might affect the offer of AS.

Urban vs. rural. No study or survey specifically addressed how urban versus rural residence might affect adherence to AS.

Academic Centers Versus Private Practice

Offer of AS. No study or survey specifically addressed whether the treatment facility's status as an academic centers versus a private practice might affect the offer of AS.

Acceptance of AS. One multivariable analysis reported that treatment facility status (academic versus community practice) was not a significant factor in predicting receiving AS/WW versus active treatment.[66]

Adherence to AS. No study or survey specifically addressed whether the treatment facility's status as an academic center versus a private practice might affect adherence to AS.

Communication Strategies

Risk Assessment, Predictive Models

No study or survey that specifically addressed the role of risk assessment and predictive models in affecting the offer of, acceptance of, or adherence to AS. One study (a 2008 review[201]) did, however, catalogue 109 prostate cancer predictive tools, which included endpoints like disease recurrence, metastasis, and survival, though no studies were identified that systematically assessed how these predictive tools were used in patient discussions.

Decision-making Tools and Aids

No study or survey specifically addressed how the use of decision-making tools or aids might affect the offer or acceptance of, or adherence to, AS.

One 2009 systematic review did, however, report on the use of various decision aids (DAs) to help men with low-risk prostate cancer participate actively in the decisionmaking process concerning their treatments.[197] Thirteen of 219 articles (representing 3 RCTs and 10 nonrandomized trials) were judged eligible for inclusion. Eligibility criteria consisted of a study population that included men with low-risk prostate cancer who had the option of RP, RT, or WW. Using the Jadad scoring system,[202] the reviewers rated two RCTs as good[203,204] and one poor.[205]

The majority of the DAs examined were developed *de novo*. They included, either alone or in combination, a written information package, consultation with a nurse or urologist, generic video, interactive computer program/CD-ROM decision aid, and a personalized multidisciplinary consultation. Most of the DAs were designed to be completed outside the clinic and after diagnosis, but prior to making a decision.

The participants in general found the DAs to be informative. One RCT reported a decrease in anxiety in participants in the intervention arm (written information package with discussion, a list of questions they could ask their physician, and an audiotape of the medical consultation) versus written information alone.[205] One RCT found that there was no difference in satisfaction with treatment choice between those who received individualized DAs and those using a generic DA.[204] One RCT found that the men in the DA arm selected their physician's treatment choice less often than those who received usual care.[203] The nonrandomized studies reported that DAs

appeared to increase patients' knowledge concerning prostate cancer and its treatments. They were also found to help encourage more active patient involvement in the decisionmaking process.

The authors noted several limitations in conducting their review, namely, too few high quality trials, heterogeneous outcome measures, and that the quality of the information provided in the DAs themselves were not assessed, which precluded determination as to whether these DAs met the quality standards set by the International Patient Decision Aids Standards Collaboration.[206]

Summary

Only two studies specifically examined men who were enrolled in an active monitoring protocol with triggers for curative treatments (as opposed to other non-AS observational management strategies).[172,173] The van As study found that the free-to-total PSA ratio and T-stage were independent predictors of time to radical treatments in patients on the protocol, while initial PSA, PSA density, Gleason score, number of positive cores, and prostate volume were not.[172] The Mills study found that decreased baseline anxiety and higher socioeconomic status were both associated with a decreased probability of willingness to consent to AS randomization (i.e., these men did not take a chance and proactively selected AS).[173]

Within the remainder of the heterogeneous studies, some tentative conclusions could be drawn concerning observational management strategies in men with prostate cancer:

- For many men, physician recommendation is an important element in helping reach a treatment decision.
- The context in which the consultation with a urologist is made (initial consultation versus second opinion visit) may be a factor in determining whether observational management strategy is offered as a treatment option or not.
- The following patient and clinical variables are potentially important in increasing the probability that a patient receives observational management strategies: increased age, presence of comorbidities, lower Gleason score, lower tumor stage, lower diagnostic PSA, membership in a lower risk group, and decreased baseline anxiety.
- The following patient and clinical variables are potentially important in increasing the probability that a patient interrupts observational management strategies to seek definitive treatments: decreased age, higher tumor stage, higher diagnostic PSA, higher PSA velocity, membership in a higher risk group, and increased anxiety.
- Physicians may have predetermined clinical notions as to when to recommend observational management strategies.
- For some men, opinions from family members and other influential people are important in reaching a treatment decision.
- Avoidance of treatment side effects is an important determinant in predicting the choice of observational management strategies.
- Prostate disease risk perceptions matter, as those who perceived prostate cancer as a more serious disease tended to choose RP over RT or observational management strategies.
- Men who are unattached (i.e., not in a permanent relationship) may have a higher probability of receiving observational management strategies versus active treatments.
- Men from lower socioeconomic strata or who are black (versus white) are more likely to receive observational management strategies.

- The type of insurance (e.g., Medicare vs. private insurance) may be a determinant in the choice of observational management strategies versus other treatments.
- Residing in an urban area (versus a rural area) may affect the probability of men receiving observational management strategies versus active treatment.
- The use of decision aids may be informative and could encourage more active patient involvement in the treatment decisionmaking process.

Key Question 4. What are the comparative short- and long-term outcomes of active surveillance versus immediate treatment with curative intent for localized prostate cancer?

In conformity with the Key Question, we systematically reviewed only comparative studies that directly compared AS (or other observational management strategies) to immediate treatment with curative intent. In order to understand the effectiveness of AS relative to active treatment options, studies of AS need a control group for comparison. As such, we did not include single-arm AS cohort studies, which cannot address comparative effectiveness questions. Thus, this review does not address evaluations of long-term clinical outcomes reported in noncomparative cohort studies, which are beyond the scope of the Key Question.

Clinical Outcomes

We did not identify any studies reporting clinical outcomes specifically of AS management strategies compared with immediate definitive treatment. No study evaluated AS where the intervention was employed in a predefined group of patients using predefined monitoring methods to identify patients who would potentially be eligible for treatment with curative intent.

Faced with a lack of studies comparing AS to immediate treatment, we elected to evaluate studies that compared other observational management strategies (largely resembling WW) with immediate treatment. One can argue that efficacy results from studies of WW may represent the low bound of the potential efficacy of AS.

In addition to previously published systematic reviews and evidence reports, our searches identified updated results from a multicenter RCT (2 publications, one on clinical outcomes and one on costs), and 12 cohort studies (2 prospective and 10 retrospective). Notably, the majority of evidence for this key question came from retrospective analyses of observational studies. Due to the differences in patient characteristics and risk profile between patients treated with observational strategies and those who received active treatments, confounding by indication is likely in these studies. Temporal changes in patient characteristics at diagnosis may affect the magnitude of the reported treatment effect sizes in observational studies. Of the studies we reviewed only two performed statistical analyses that accounted for year of diagnosis.[207,208] Another study included only patients diagnosed in a single year.[159] One study used pathology information based on a contemporary reading of available pathology material in its statistical adjustments.[209]

Findings from Previous Systematic Reviews

We examined two recent systematic reviews of treatments of men with clinically localized prostate cancer.[8,9] Of note, these systematic reviews included some studies of men receiving

ADT in their observational management groups, such primary studies would not have met the primary study inclusion criteria for this report.

One AHRQ evidence report[9] included two RCTs comparing WW with RP: the Scandinavian Prostate Cancer Group Study 4 (SPCG-4),[210] and the Veterans Administration Cooperative Urological Research Group (VACURG) trial.[211] A recent (2010) Cochrane Report on the same topic[212] did not identify any additional studies; however, we identified the latest update of the SPCG-4 trial results[213,214] (discussed in the primary study section below). In SPCG-4, 695 patients were enrolled between 1989 and 1999, and randomized to either watchful waiting (WW) or radical prostatectomy (RP); they were followed for a median of 8.2 years. When compared with patients on WW, patients who had RP had significantly lower mortality (RR 0.74; 95 percent CI 0.56, 0.99; P = 0.04), disease-specific mortality (RR 0.56; 95 percent CI 0.36, 0.88; P = 0.01), and distant metastases (RR 0.60; 95 percent CI 0.42, 0.86; P = 0.04).[210] The VACURG trial followed 142 patients for a median of 23 years, and found no difference in mortality between WW and RP groups.[211]

Additionally, the AHRQ evidence report on localized prostate cancer treatment[9] considered the results of two randomized trials reporting on QoL and self-reported functional status.[215,216] One included study was an ancillary investigation from the SPCG-4.[215] The study found that self-reported erectile dysfunction and urinary leakage were more common in the RP group, whereas urinary obstruction was more common in the WW group. Bowel function, prevalence of anxiety, prevalence of depression, well-being, and the subjective QoL were similar in the two groups.[215] The second study was based on a randomized trial comparing RT with deferred treatment[216] and demonstrated that patients in the RT group experienced a decrease in QoL due to the development of hematuria, incontinence, mucus, and having to plan daily activities in response to intestinal problems.[216]

Another AHRQ report[8] that compared RT and no treatment evaluated one prospective cohort study[217] and eight[209,218-224] retrospective cohort studies. The included prospective cohort study reported no difference in sexual function between brachytherapy (BT) and no treatment, but significantly worse sexual function between external beam radiation therapy (EBRT) and no treatment.[217] Of the four retrospective cohort studies that compared disease-specific survival between radiation therapy and no treatment, one found significantly better disease-specific survival in men treated with BT.[219] Three studies reported gastrointestinal or genitourinary toxicity outcomes and found no difference between BT or EBRT and no treatment, but one study found a higher rate of receiving treatment for urethral stricture in patients treated with combined EBRT and BT, compared with those with no treatment.[218] One study reported significantly higher rates of second primary cancer in patients treated with EBRT compared with those with no treatment.[220]

Findings from Primary Studies

To address Key Question 4, we searched for studies that compared observational management strategies with immediate definitive treatment. As discussed above, our searches did not identify any studies comparing patients managed with AS versus immediate active treatment with curative intent. Here we summarize findings from reviewed studies where the observational management strategy mostly resembled WW. We included only multicenter studies that enrolled men with localized prostate cancer, and reported age-adjusted effect sizes. Characteristics of the 12 eligible studies reporting on clinical outcomes are shown in Appendix Table C4.1. Two RCTs (in 3 publications[213,214,225]) three prospective cohort studies,[226-228] and ten

retrospective cohort studies[207-209,218,223,224,229-232] were included. The two RCTs reported updated results from the SPCG-4 and UMEA1 trials that had been included in the previously mentioned systematic review.[9] Among the 13 cohort studies, sample size ranged from 113[226] to 44,630[208] and followup duration from 12 months[226] to more than 13 years.[209] Methodological quality of the studies was rated as A in one study,[213] B in nine studies,[207-209,214,218,223,224,229,230] and C in five studies.[226-228,231,232]

Comparison Between Observational Management Strategies and Radical Prostatectomy[d]

Appendix Table 4.2 presents the detailed results from the one RCT and its ancillary investigation, as well as 14 cohort studies that compared observational management strategies with RP.

Randomized controlled trials. The SPCG-4 RCT followed 695 men with localized prostate cancer for a median of 12.8 years.[213] Compared with men on WW, men treated with RP had significantly lower prostate cancer-specific mortality (RR 0.62; 95 percent CI 0.44, 0.87; P = 0.01), all-cause mortality (RR 0.75; 95 percent CI 0.61, 0.92; P = 0.007), and incidence of distant metastases (RR 0.59; 95 percent CI 0.45, 0.79; P <0.001).

Subgroup analyses found no significant modification of the treatment effect on mortality by PSA (< 10 vs. ≥ 10 ng/mL, interaction P = 0.72 and P = 0.30, for overall and prostate cancer-specific mortality, respectively) or Gleason score (<7 vs. ≥7, interaction P = 0.36 and P = 0.52, for overall and prostate cancer-specific mortality, respectively). However, age was found to be a modifier of the treatment effect (interaction P = 0.003 when age was dichotomized at 65 years of age; P = 0.001 when treating age as a continuous variable). The favorable effects of RP on overall mortality were present among men younger than 65 years (HR = 0.52; 95 percent CI 0.37, 0.73; P < 0.001), but not men older than 65 years (HR = 0.98; 95 percent CI 0.75, 1.28; P = 0.89). Effect modification by age did not reach statistical significance for prostate cancer-specific mortality (interaction P = 0.16; HR = 0.49; 95 percent CI 0.31, 0.79; and HR = 0.76; 95 percent CI, 0.25, 2.32; comparing men younger vs. older than 65 years, respectively). The authors reported that none of the subgroup analyses performed were specified in the main study protocol but were determined "before any data were seen."[213] Thus, the study may not have had adequate power to detect effect modification; the absence of statistically significant interactions does not indicate that clinically meaningful differences do not exist between the investigated subgroups.

An additional subgroup analysis limited to low-risk prostate cancer patients (124 in the RP group and 139 in the WW group) was also reported. Low-risk cancers were defined as PSA < 10 ng/mL and Gleason score < 7 or WHO grade of 1 (for tumors diagnosed only by cytologic assessment). In this subgroup, the HRs comparing RP versus WW were 0.62 (95 percent CI 0.42, 0.92; P=0.02), 0.53 (95 percent CI 0.24, 1.14; P=0.14) and 0.43 (95 percent CI 0.23, 0.79; P=0.008), for overall survival, prostate-cancer specific mortality and development of metastatic disease, respectively. The study reported that the most common non-fatal complications in the RP group within 1year of surgery were impotence and urinary leakage, with cumulative one-year incidence proportions of 58.1 percent and 32.2 percent, respectively. Urinary obstruction,

[d] During the conduct of our review, preliminary results from the Prostate cancer Intervention Versus Observation Trial (PIVOT, NCT00007644), a randomized trial of RP versus WW with palliative intervention, were presented at the 2011 annual meeting of the American Urological Association (AUA). We did not include this study in our evidence tables because no full text publication was available.

pulmonary embolism and deep-vein thrombosis were uncommon (cumulative one-year incidence proportions 2.1 percent, 1.4 percent and 1 percent, respectively). Data on adverse events in the WW group for the first year were not provided.

An ancillary investigation[214] from the Swedish component of the SPCG-4 RCT (including a total of 326 patients) evaluated the quality of life outcomes at different followup timepoints (2 to 3 years, 4 to 5 years, and 6 to 8 years postrandomization). The study found that anxiety (moderate or high) and depressed mood (moderate or high) were less common in the RP group compared to the WW group, and that sense of well being and quality of life were not significantly different between groups at 2-3 years post randomization. Depressed mood was more common in the RP group compared to the WW group at 4-5 years post randomization; significant differences were not observed for other quality of life outcomes. No significant differences were observed at and 6-8 years post randomization. This ancillary investigation also reported that about 80 percent and 42 percent of patients in the RP group reported erectile dysfunction and urinary leakage, compared to 37 percent and 11 percent in the WW group respectively 2 to 3 years post randomization. At 6-8 years post randomization the corresponding frequencies were 83 percent and 56 percent in the RP group and 55 percent and 25 percent in the WW group.

Observational studies—Cancer-specific mortality. Three cohort studies[207,209,223] compared prostate cancer-specific mortality between patients on observational strategies and patients treated with RP based on large databases (SEER-Medicare,[207] the National Cancer Register of Sweden Follow-up Study, NCRSFS,[223]and the Connecticut Tumor Registry[209]). All three studies identified a statistically significant difference between treatments favoring RP (HR 0.63; 95 percent CI 0.55, 0.71 [using a propensity score[207]], HR 0.49; 95 percent CI 0.34, 0.71,[223] and HR 0.29; 95 percent CI 0.17, 0.52,[209]respectively). However, an instrumental variable analysis of the SEER-Medicare study did not identify any significant difference between treatments with wide confidence intervals around the estimated treatment effect (HR 1.37; 95 percent CI 0.15, 12.5).[207]

Observational studies—All-cause mortality. Five studies compared all-cause mortality between observational management strategies and RP.[207,208,223,230,232] One was based on the CDC-NPCR Patterns of Care Study (POCS),[230] two used the SEER-Medicare database,[207,208] one used NCRSFS data,[223] and one used the Center for Prostate Disease Research database to identify men of 70 years of age or older who were diagnosed with localized prostate cancer.[232] From CDC-NPCR POCS, the 5-year all-cause mortality rate was significantly higher in patients on WW than in patients treated with RP (adjusted HR 2.3; 95 percent CI 1.70, 3.12).[230] Similarly, the two reports that analyzed data from the SEER-Medicare database,[207,208] one report analyzed data from the Center for Prostate Disease Research database,[232] and the NCRSFS study[223] showed favorable outcome in the RP group (HR 0.65; 95 percent CI 0.62 ,0.68[using a propensity score[207]]; adjusted HR 0.49; 95 percent CI 0.41, 0.57[207]; adjusted HR 0.50; 95 percent CI 0.47, 0.53;[208] and adjusted HR 0.52; 95 percent CI 0.32, 0.84,[232] respectively; the latter estimate was derived from an analysis comparing WW to RP (reference group), where patients initially managed with WW were stratified by whether they later received secondary therapy. This analysis may be susceptible to bias because the decision to administer secondary treatment is often made based on information not available at baseline (e.g., tumor progression status during followup); and this information is likely to be associated with the clinical outcome of interest. This study was exclusively limited to patients who were aged 70 years or older.[232] An

instrumental variable analysis performed in one of these studies however, did not find a significant difference between treatments (HR 0.92; 95 percent CI 0.39, 2.17); however confidence intervals were wide indicating substantial uncertainty around the HR estimate.[207]

Observational studies—Morbidity of primary treatment. One report analyzed the data from the CaPSURE database and found that men treated with RP had a higher rate of receiving treatments for urethral stricture than men on WW over a median followup of 2.7 years (adjusted HR 10.4; 95 percent CI 3.28, 33.3).[218] One study analyzed the risk of additional surgical procedures following primary treatment with WW, RP, RT or ADT.[231] The reported multivariable-adjusted estimates compared WW only with RP and demonstrated that bladder irrigation/cystostomy procedures and TURP/bladder neck incision were more common in the WW group compared to the RP group (HR 1.71; 95 percent CI 1.33-2.20 and HR 2.63; 95 percent CI 2.08-3.33, respectively). In contrast urethra dilation procedures were less common in the WW group (HR 0.71; 95 percent CI, 0.61-0.84) and cystoscopy procedures were not significantly different (HR 1.00; 95 percent CI, 0.88-1.13), compared to the RP group.

Observational studies—Quality of life. Four studies reported QoL outcomes.[226-229] Two studies analyzed data from the CaPSURE registry. One of them reported that patients on observational management strategies (WW or AS, reported in aggregate) had significantly lower mean SF-36 scores in the social function domain compared to patients treated with RP (89 vs. 100; $P < 0.05$, respectively), but no difference was found in three other domains of SF-36.[229] The other study using CaPSURE data[227] presented trends in QoL scores from immediately after treatment through 2 years of followup, and reported that scores for most general and disease-specific QoL domains improved statistically significantly at one and two years post treatment compared with scores immediately after treatment among patients who received RP. Patients on observational management strategies (WW or AS, reported in aggregate) did not experience statistically significant changes in their scores over time with the exception of a statistically significant improvement in "health change" at one year of followup (compared to immediately post treatment assignment), however the sample size of the group was small (n=87). Patients receiving radiotherapy experienced improvements in "health change," bowel function, and urinary bother scores between the immediate post-treatment period and one year of followup (however, "health change" significantly worsened between year 1 and year 2 in this group).

The third study originated from four academic centers in Wisconsin and suggested that disease-specific QoL declined more among patients receiving RP compared to WW (for the domains of urinary and sexual function) but did not identify a difference between treatment groups for urinary bother, sexual bother, bowel function, bowel bother, or general QoL.[226]

The fourth study included a subgroup of patients with available longitudinal data from the Health Professionals Followup Study followed for up to 3 years to assess the changes in general QoL and UCLA Prostate Cancer Index. The study reported that WW was not associated with statistically significant changes on any of the SF-36 scales compared to RP (data were extracted only for the longitudinal component of this study).[228]

Studies of QoL often report multiple outcomes and perform multiple comparisons. None of the studies we reviewed reported the use of statistical procedures that account for multiple testing. However, we note that such procedures are not universally recommended.[233,234]

Comparison Between Observational Management Strategies and Radiation Therapy

Detailed results from the eight studies that compared observational management strategies and radiation therapy (RT) are shown in Appendix Table C4.3. No study reported on the effects of treatment on the development of metastatic disease.

Randomized controlled trial. One study[225] reported quality of life comparisons based on a subset of 54 surviving responders from the UMEA1 trial. UMEA1 was a randomized trial of 166 patients comparing RT with WW. The authors stated that the complete trial data are "not yet ready for publication." At 10 years of followup, the study reported no statistically significant differences in health-related quality of life function or symptom scales. No significant difference was observed between RT (mean=2.3) and WW (mean = 2.3; P not reported) for "limitations in daily life" caused by prostate cancer. No significant difference was observed for "life situation" between RT (mean=3.7) and WW (mean=3.3; P=0.398). Out of 20 comparisons in symptom/function scores (abstracted in Appendix Table C4.3) only that of "sexual bother" was statistically significant (mean scores: 3.8 vs. 7.4 in the WW and RT groups, respectively; higher values indicate worse function; P=0.011). A weak urinary stream was the most reported symptom and the mean score differed significantly between WW (mean 4.8) and RT (mean 3.0; higher values indicate worse function; P = 0.034). No differences were observed at 10 years between RT and WW regarding the questions "How would you feel if you lived the rest of your life with your urinary/bowel/sexual problems as they are now?" (P values comparing the two groups were 0.643, 0.653 and 0.819, respectively).

Observational studies—Cancer-specific mortality. One study using data from NCRSFS reported that RT was associated with a non-significant improvement in prostate cancer-specific mortality compared to surveillance (WW or AS in aggregate; HR 0.70; 95 percent CI 0.45, 1.09) at 8.2 years of followup.

Observational studies—All-cause mortality. Two studies reported adjusted estimates comparing all-cause mortality between observational management strategies and RT.[208,223] One study was based on the SEER-Medicare database,[208] and the other used NCRSFS data.[223] Both studies reported significantly lower mortality rates in patients treated with RT compared with patients managed with an observational strategy (HR 0.81; 95 percent CI 0.78-0.85,[208] and HR 0.68; 95 percent CI 0.57, 0.82,[223] respectively).

Observational studies—Morbidity of primary treatment. One report analyzed the incidence of treatment for urethral stricture captured in the CaPSURE registry and did not find a significant difference between patients on WW and patients treated with EBRT, or between patients on WW and patients treated with BT over a median followup of 2.7 years.[218]

Observational studies—Quality of life. Three studies reported quality of life outcomes.[224,226,229] From the CaPSURE registry, patients on WW had significantly lower mean SF-36 score in the social function domain than patients treated with RT (89 vs. 86; P < 0.05, respectively), but no difference was found in three other domains of SF-36.[229] In a study of 4 academic medical centers from Wisconsin no significant difference in disease-specific and general QoL was observed between patients managed with RT and those managed with expectant management.[226] Finally, a study based on the Eindhoven Cancer Registry found that RT had a negative effect on

the physical functioning and bodily pain dimensions of the SF-36 instrument, the spiritual and total wellbeing scores of the Quality of Life-Cancer Survivors (QoL-CS) instrument, and the bowel function and bowel bother dimensions of Expanded Prostate Cancer Index Composite (EPIC), compared to observational management. No other significant difference between treatments was observed for general QoL, cancer-specific QoL, or disease-specific-QoL.

Comparison Between Observational Management Strategies and Combination Therapy or Active Treatments Considered in Aggregate

Detailed results from the two studies that compared AS/WW and other active treatment groups (a combined group of patients managed with RP or RT and a group receiving both RT and EBRT) are shown in Appendix Table C4.4. No study reported incidence of metastatic disease or quality of life outcomes.

One study analyzed data from the SEER-Medicare database up to year 2002, and compared patients on observation with patients who received any active treatment, including RP, EBRT, and BT.[208] Compared with patients on observation, patients on active treatment had significantly lower risk of prostate cancer-specific mortality (adjusted HR 0.67; 95 percent CI 0.58, 0.77) and all-cause mortality (adjusted HR 0.69; 95 percent CI 0.66, 0.72).[208]

One study analyzed the data from the CaPSURE registry and found that a group of men treated with EBRT and BT (combined treatment) had a higher rate of receiving treatments for urethral stricture than men on WW over a median followup of 2.7 years (adjusted HR 4.56; 95 percent CI 1.23, 16.88).[218] No significant difference was found between patients on WW and patients treated with combined RP and EBRT.[218]

Costs

Findings from Studies of Actual Patient Costs

We identified four primary studies (three using U.S. data[159,235,236] and one from Sweden[237]) reporting on comparisons of costs of active treatments and observational management strategies for localized prostate cancer. All studies included groups of patients treated with WW and used various active treatments as comparators; no comparative study reporting on costs included a group managed with AS. Details from each study including the populations, treatments compared, and cost estimates are presented in Appendix Table C4.5.

One study used the SEER-Medicare database (13,769 patients matched 1:1 with noncancer controls; 2805 of the patients had been managed with WW) to estimate incremental treatment costs during the first 5 years of treatment. Using inverse probability of treatment weights derived from a propensity score to account for factors that affect treatment selection, this study found that, WW has lower incremental costs ($8535) compared with RP ($19,481) or RT ($16,653) over 5 years.[159]

A second study used data from the CaPSURE database (235 patients; 37 managed with WW) to estimate mean first year costs.[235] The unadjusted mean cost of WW ($484) was lower compared with RP (without hormonal therapy, $7320) or RT (without hormonal therapy, $7430). After adjusting for patient and disease characteristics, the difference in costs among treatments was statistically significant (analysis of covariance $P < 0.001$).

A third study, again using CaPSURE data, estimated that WW had a total cumulative cost over 5.5 years of follow up of $31,871 and $31,789, for low and intermediate risk prostate cancer patients, respectively. The corresponding values were $28,366 and $41,419 for BT; $48,840 and

$56,725 for EBRT and $32,795 and $35,037 for RP.[236] Notably, each treatment had a distinct cost pattern when total costs were separated into those related to medications, office visits and hospitalizations (please see Appendix Table 4.5 for details).

The fourth study was an ancillary investigation from the SPCG-4 trial, a randomized study of RP versus WW for men with localized prostate cancer.[237] The study reported cost estimates (based on Swedish prices and converted to euros, €) for a subset of patients (n=212; 105 managed with WW and 107 with RP) participating in the trial. After a median followup of 11.8 years for the WW group and 12.2 years for the RP group, the total mean cost of WW was €18,124 for WW compared with €24,147 for RP. After adjustment for age, Gleason score and PSA, the difference between treatments remained and was statistically significant (P = 0.003). The applicability of Swedish cost estimates to the United States setting is likely to be limited.

Findings From Model-Based Economic Evaluations

In addition to the primary cost studies discussed above, we also considered model-based cost estimates, including the cost-minimization or cost-consequence analyses and cost components of cost-effectiveness or cost-utility analyses of AS (but not WW). We did not review the cost-effectiveness or cost-utility analyses reported in these documents and focused only on cost information.

Two economic evaluations conducted by the Institute for Clinical and Economic Reviews (ICER)[238,239] reported cost-utility analyses comparing AS with active treatment options (IMRT, BT, and open retropubic nerve-sparing RP). [e] An associated publication based on these reports was also reviewed.[240] Cost estimates were obtained from multiple sources including outpatient costs from the 2008 Red Book, Hospital Outpatient Prospective Payment System, Physician Fee Schedule, Centers for Medicare & Medicaid Services Lab Fees and Durable Medical Equipment Schedules; and inpatient payments from the Hospital Inpatient Prospective Payment System, the 2008 Anesthesia Conversion Factor and American Society of Anesthesiologists payment information. The information obtained was incorporated into a disease history Markov model to determine the cost (and effectiveness) of each treatment strategy. The AS protocol included regular physical examinations, PSA testing, and rebiopsies (one year following diagnosis and every 3 years thereafter). In the model, 61.1 percent of all patients originally assigned to AS received treatment, using a lifetime horizon; 28.3 percent of all patients received treatment within the first 5 years. In the base case analysis (cohort of 65 year old men diagnosed with low-risk clinically localized prostate cancer, lifetime model horizon, 3 percent annual discount rate) comparing AS (followed by IMRT with short term ADT in cases of progression, or IMRT alone in cases of patient preference for active treatment) and RP, the total cost of an open RP management strategy was estimated at $28,348 and the total cost of an AS strategy was estimated at $30,422. In a similar model (cohort of 65 year old men diagnosed with low-risk clinically localized prostate cancer, additional life expectancy of 16 years, lifetime model horizon, 3 percent annual discount rate) the total costs were $30,422 for AS, $23,348 for RP, $25,484 for BT, $37,861 for intensity-modulated RT and $53,828 for proton beam RT. Sensitivity analyses produces largely consistent results.

One study[241] reported a cost comparison of AS and RP. Costs were obtained from 2008 Medicare reimbursement rates. The costs for RP were based on Medicare reimbursements in 2008 for patients treated by a single, high-volume surgeon (volume was not reported). Several AS protocols were modeled: they differed by biopsy frequency (within a year of diagnosis and

[e] ICER reports are available online at http://www.icer-review.org; last accessed September 30, 2011.

every 12 or 18 months thereafter), duration of active followup during which biopsies were performed (5 or 10 years after diagnosis) and the rate of "conversion" to active treatment (5 percent or 7 percent annually, obtained from literature sources). The model had a 15 year horizon and used a 5 percent discount rate. The cost per patient in the AS models ranged from $6558 to $11,992. Costs were higher with longer periods of active followup, more frequent biopsies and higher rates of conversion to active treatment. Costs for the AS strategies were 43 percent to 78.7 percent of the per-patient cost of RP. In the RP group 92 percent of all costs occurred in the first year; in contrast, most costs in AS strategies occurred in later years. Sensitivity analyses demonstrated that increasing the cost of RP by $5000 affected the cost comparison results: it led to a 14.2 percent to 20.2 percent increase in the cost of AS versus a 32.8 percent increase in the cost or RP.

Another study[242] reported a cost comparison of AS versus open retropubic RP, robotic-assisted RP, EBRT, and BT. Professional fees were obtained from 2010 Medicare reimbursement values for the study region (Miami, Florida). For inpatient costs the model used the mean inpatient costs of patients with clinically localized prostate cancer at the authors' institution (Miller School of Medicine, Miami, Florida). The AS protocol included transrectal ultrasound guided biopsies (within a year of diagnosis and annually thereafter), follow-up visits for PSA testing and digital rectal examination of the prostate (every 3 months for the first 2 years and every 6 months thereafter). The estimates also incorporated a 25 percent probability of receiving treatment at 5 years on AS (10 percent were expected to receive EBRT and 15 percent open retropubic RP). A disease history Markov model was used to determine the cost of management of complication associated with different treatment strategies. The incidence of the various complications was obtained from Hayes et al.[240] The model horizon was 10 years and the reported stated that the model did "not account for inflation." At year 1, the cumulative cost was $9732 for retropubic RP, $17,824 for robotic-assisted RP, $20,730 for EBRT, $14,061 for BT, and $1154 for AS. At year 10, the cumulative cost was $15,084 for retropubic RP, $22,762 for robotic-assisted RP, $23,953 for EBRT, $17,284 for BT, and $13,116 for AS. The authors noted that AS is associated with a different cost distribution because its initial cost is lower but the cost of follow-up is relatively higher, compared to active treatments.

We caution that the cost estimates discussed in this section are model-based; as such, they are sensitive to the model structure and inputs, including model assumptions about disease natural history and treatment effectiveness.

Summary

No study reported clinical outcomes specifically for AS management strategies versus immediate definitive treatment. Therefore, there is insufficient evidence for the comparative short- and long-term outcomes of AS versus immediate definitive treatment for localized prostate cancer.

We identified an updated analysis from a multicenter RCT (SPCG-4) and 11 multicenter cohort studies that reported clinical outcomes comparing observational management strategies with active treatments including RP and RT. We also identified a cost study based on the above mentioned SPCG-4 trial and two additional observational studies comparing the costs of treatments for localized prostate cancer. The majority of evidence for Key Question 4 came from observational studies. Confounding bias (often referred to as "confounding by indication") is a concern for such studies, due to the differences in patient characteristics (that may be associated with the outcomes of interest) between patients treated with observational strategies and those

who received active treatments. Although multivariable regression analyses or propensity score methods were employed to control for confounding by all reviewed studies, such analyses cannot account for unmeasured confounders of the treatment-outcome association.

Observational management strategies versus RP. Studies generally reported that men treated with RP had lower all-cause or prostate cancer-specific mortality rates than men on WW. The development of metastatic disease was assessed by a single study that found a significant benefit for RP compared to WW. Morbidity of primary treatment was reported by two studies that suggested an increased risk of urethral stricture (and procedures to treat it) were less likely among patients managed using observational management strategies. One of these studies also investigated cystoscopies (equally common in RP and observation groups), bladder irrigation/cystostomy and TURP/bladder neck incision (both more common among patients managed with observation). QoL was reported in three studies, which reported heterogeneous results.

Observational management strategies versus RT. Studies generally reported that men treated with RT had lower all-cause mortality rates than men on WW. One study reported prostate cancer-specific mortality information and found no statistically significant difference between RT and observational management. No study reported on treatment comparisons for the development of metastatic disease. Morbidity of treatment decision was reported by only one study which found no significant difference between observational management and BT or EBRT. QoL measures were reported in four studies, which reported heterogeneous results.

Observational management strategies versus combined radiation modalities or active treatments considered in aggregate. Data from one study showed that active treatments (RP, RT, BT considered together) resulted in lower all-cause and prostate cancer-specific mortality rates compared to WW. Morbidity of primary treatment was reported by only one study which found that a group of patients receiving EBRT and BT (combination therapy) had a higher rate of receiving treatments for urethral stricture compared to a group managed observationally.

Short- and long-term costs. Both short and long-term costs observed in clinical studies appear to be higher for active treatment strategies (RP or RT) compared to WW; however evidence originated from small studies using heterogeneous measurement methods. We did not identify any primary study comparing the cost of AS with active treatment strategies; economic modeling using U.S. prices suggested that within 10 to 15 years of diagnosis AS may be less costly compared to active treatments; a study using a lifetime horizon indicated that AS may be associated with higher costs compared to RP and BT, but lower costs compared to intensity modulated RT (IMRT) and proton beam RT.

Key Question 5. What are the research needs regarding active surveillance (or watchful waiting) in localized prostate cancer?

As summarized and discussed in previous sections of this report, the evidence directly addressing the four principal Key Questions is largely incomplete. In part this is because published studies tended to address research questions that were different in scope or focus than the questions posed by the sponsors of the NIH State-of-the-Science Conference; in part much of the available data are not amenable to analyses that could adequately answer the Key Questions.

As described for Key Question 2, there is not yet consistency among clinicians or researchers as to the definitions of AS or of WW, the standard protocols for the interventions, or how to manage patients whose cancers show signs of progression. In addition, studies based on database analyses frequently could not segregate patients who had AS from those who had WW. It is also common for analyses to group together patients who had no treatment with those who had ADT alone. This was particularly the case for analyses of the SEER database, in which it is only possible to distinguish nonaggressive therapy (e.g., AS or ADT) from prostatectomy or radiation therapy.

For the most part, the Key Questions, and thus this review, implicitly assume that it is possible to identify men who are at sufficiently low risk of progression of their prostate cancer that AS can be a safe and appropriate option for them. This report did not review studies that evaluated methods of categorizing patients into risk groups. However, our readings and discussions with experts lead us to the conclusion that additional basic and clinical research is needed to more accurately classify or predict those men whose diseases are indeed at a low risk of progression. Presumably, these are the men for whom it would be most appropriate to consider offering AS.

Also as described above, there are numerous gaps in the evidence regarding the many specific factors and subgroups of interest to the conference. This section will not attempt to delineate all the places where evidence is inadequate, but instead will highlight those areas that our EPC concluded are in most need of future research. The future research needs will be addressed in the order of the Key Questions.

Key Question 1. Patient Population and Natural History Changes in Last 30 Years

While there are several gaps in evidence regarding time-trend analyses of specific factors of interest to the conference sponsors, better understanding of time-trends in the future can be gained by improving the data being collected and expanding the scope of the major U.S. databases. In particular, we found that detailed stage and grade information are often not available requiring researchers to create broad categories that place major limitations on the analyses. Likewise, future research would be enhanced if more accurate and specific data were collected about the therapeutic interventions or observational management strategies. As just mentioned above, one cannot use the SEER database to accurate analyze true WW (or AS), since these observational management strategies cannot be distinguished from use of ADT. In addition, the SEER database is inadequate to analyze data from races other then blacks and whites, since Hispanics/Latinos, Asians, and others are apparently underrepresented, precluding complete racial analyses. This may require adding new registries to SEER that better represent other races.

We were also concerned about a potentially important source of bias in the SEER database which may require resolution to allow for appropriate future analyses on cancer staging. Analyses of SEER reported summary stage information using the "a combination of the most precise clinical and pathological documentation of the extent of disease."[f] This may result in misclassification, as the accuracy of available information on staging depends on the treatment patients receive. For patients who have had RP, pathological staging information is available, whereas for those who have received RT, ADT monotherapy, AS, or WW, only clinical staging

[f] See the SEER staging manual, available at: http://seer.cancer.gov/tools/ssm/ (last accessed September 30, 2011).

information is available. Thus, patients having surgery are staged more accurately than those with clinical or imaging staging alone. This bias could be reduced if the SEER database maintained the staging information that is available prior to surgery, so that researchers can analyze unbiased data about staging.

Key Question 2. Definition of Active Surveillance

Little new research per se is needed to address how AS has been defined by researchers. However, interpretation of future studies would be best served if there were a standard, agreed-upon definition of AS that clearly distinguishes it from WW and other forms of withheld or noncurative treatments. A consensus conference may be the most appropriate forum to define AS. Features of the definition will need to include 1) the goal or intent of the intervention (e.g., delaying curative treatment until there is evidence of progression); 2) the "eligibility criteria," a determination of which patients should be offered AS based on disease and patient characteristics; 3) the "followup protocol," the minimum set of tests that should be followed (e.g., DRE and PSA), and their timing; and 4) criteria or triggers for stopping AS, when there has been sufficient or rapid enough progression to warrant active, curative treatment.

Working under the (still unproven) assumption that AS is a safe and effective treatment alternative to RP or RT, the best AS protocol should be investigated by randomized or other prospective comparative studies that directly compare different protocols. The current retrospective or case series studies provide some data to allow for comparison of protocols, but these data are largely incomplete and adjustment using techniques of multivariable analysis is likely inadequate to control for confounding and other biases. Examples of comparisons for future trials could include use of different combinations of followup testing (e.g., PSA, DRE, imaging, rebiopsy), different timing for the tests (e.g., every 3 or 6 months), and different definitions of progression that would determine when curative treatment is offered. These trials will require long-term followup. The outcomes of greatest clinical importance are those that are most pertinent to patients' health, well-being, and longevity. Examples include all-cause mortality, prostate-cancer-specific mortality, development of symptomatic disease or distant metastases, urological and other complications (from testing, treatment, or deferring treatment), quality of life, anxiety, and family dynamics. Also of interest would be overall costs, use of resources, and numbers of negative invasive tests (i.e., biopsies showing no progression that arguably were thus unnecessary). Since only about half of men on AS require treatment due to disease progression within 5 years,[168] and only a percentage of them will have clinically important outcomes (e.g., cancer death), a trial may also need to be quite large to be adequately powered.

Another related question of interest that was not asked for in the Key Question (and thus was not systematically reviewed) is which tests are the best predictors of either progression or clinical outcomes. Ideally, for the purpose of developing an AS protocol, these studies should be conducted only in men who are being followed with AS, excluding men with more advanced disease at baseline or who are undergoing curative treatment. Studies would need to properly account for whether ADT is being used. Prospective studies that directly compare specific tests (e.g., PSA, DRE, imaging, rebiopsy) would be most reliable. However, such studies may have large amounts of confounding and colinearity. For example, there will likely be large variation in the frequency of specific tests, which may be confounded with the tests themselves; the results of some tests (e.g., DRE) may affect the frequency or use of other tests (e.g., rebiopsy), and it may

be impossible to separate out the effects of individual tests that are conducted together (e.g., PSA and DRE).

Consideration should be given to conceptualizing AS monitoring strategies as dynamic treatment regimes (i.e., rules for sequential decision making based on the evolution of patient or tumor characteristics over time). Such approaches formalize the process of choosing between competing monitoring strategies based on expected responses to treatment and related intermediate and long-term outcomes using appropriate statistical models. Compared to standard research methods (e.g., directly comparing two monitoring strategies in a parallel group study), dynamic treatment modeling may be better at identifying the optimal monitoring regime while accounting for the temporal structure of the data (e.g., multiple monitoring visits) and the fact that treatment decisions at each visit are determined by the measurements performed (e.g., PSA, repeat biopsy). Indeed, statistical methods exist that can use observational or randomized study data to determine the factors that should be considered as triggers for intervention, as well as the optimal cut-off values of these factors.[243,244]

At a minimum, future study reports should be very explicit and clear about what their definitions of AS (or WW) were, what were the goals of the intervention, what were the exact protocols, what were the exact definitions of progression, how and when protocols or standards changed during their study (and why), and why and how often patients and clinicians chose to not follow the protocols.

Key Question 3. Factors That Affect Offer of, Acceptance of, and Adherence to AS

As described under the findings for Key Question 3, there are two major categories of studies that address this Key Question: quantitative analyses of databases and registries, and more qualitative analyses of surveys of men diagnosed with prostate cancer and their clinicians. To date, both types of analyses have limitations that preclude strong conclusions. The databases tend to have data only about what treatment patients received and when. Therefore, whether different treatment options were offered to them, whether they accepted those options, and whether they adhered to their initial choices could only be inferred. Even the best analysis of predictors of initial treatment cannot adequately address the Key Question of interest to this conference's sponsors, since the three treatment stages of interest (offer, acceptance, and adherence) are not described in the database. Thus, full statistical analyses of predictors will require the prospective collection of data specifically about what interventions were offered to each patient, which treatments the patients accepted, and when they chose to receive curative treatment despite lack of evidence of progression. Ideally, data would also be collected on what *a priori* definition of progression was used for each patient to allow the analysis of lack of adherence. These datasets will need to be sufficiently large to allow for testing of multiple predictor variables. In addition, future studies should only perform complete analyses of all treatment options (AS or WW, surgery, radiation, ADT, and combinations) without arbitrarily grouping treatments (e.g., AS and ADT) or selectively excluding treatments (e.g., by pairwise comparisons). This will minimize bias and increase clarity about what is being tested.

We believe that, where possible, future database analyses and prospective observational studies should focus on those predictors that are amenable to change or that can be acted upon. For example, if it is shown that men who receive educational materials are more likely to accept AS, this intervention can be implemented. Or if it is found that black men are less likely to be offered AS, then training of physicians to minimize implicit bias may be warranted. However,

researchers should avoid interpreting analyses to suggest that men with certain demographic (or other nonmodifiable) features are most likely to accept treatment and thus other men should not receive the offer of treatment.

Further surveys of patients, their families, and their clinicians are warranted. To improve reliability, these should be adequately powered to ensure that sufficient numbers of men were treated with different interventions and to allow full analyses of the tested predictors. Studies should use established methods including standardized qualitative research designs and, ideally, validated questionnaires to elicit preferences. Studies of this sort also need to consider the overall adequacy of discussion with patients regarding management options—and documentation of those discussions. Adequate documentation of these discussions will surely improve the veracity of some of these survey data.

Though not requested by the sponsors, future Key Questions of interest to be addressed by systematic review and primary studies could include comparisons of interventions that improve the likelihood that eligible men are offered AS, that improve acceptance of AS, and that improve adherence with AS. Arguably, it is more important to first establish how to successfully get men offered, accepting, and adhering to AS before determining which men are at greatest risk of failing to receive AS. If no intervention successfully improves the likelihood that men will adhere to AS, it may not be particularly relevant to flag those men most at risk of nonadherence.

Another issue for consideration could be when and how to discuss with patients the option of transitioning from AS to either WW or other nontreatment protocols, for those patients who may decide that they might no longer desire curative treatment regardless of progression.

Key Question 4. Active Surveillance Versus Immediate Curative Treatment

An RCT with long followup would provide the best evidence to adequately assess the differential effects between AS and immediate curative treatment. The least biased, most reliable study design comparing two interventions is the RCT that adheres to modern methodological standards. While the patient and his clinicians cannot be blinded to his treatment plan, outcome assessors—particularly those who conduct psychometric testing—should be blinded. The primary outcomes of interest should be the same as those listed above, under research needs for Key Question 2, namely patient-centered clinical outcomes, including psychometric measurements, adverse events, resource utilization, and costs. However, we acknowledge that conducting and completing an adequately powered trial of sufficient duration may be challenging and resource-intensive. The greatest difficulty is likely to be recruiting sufficient physicians and patients who are willing to allow chance to dictate the choice between AS and immediate treatment (note that the Active Surveillance Therapy Against Radical Treatment in Patients Diagnosed With Favorable Risk Prostate Cancer (START) has been terminated because of not meeting accrual target[g]). One will also need to factor in the probability of acceptance of the scientific equipoise across these treatment options in maximizing recruitment (witness the rapid enrollment in the Prostate Testing for Cancer and Treatment trial (ProtecT) discussed further in the Discussion). Trials would then need to be of sufficiently long duration to collect data on the clinically relevant outcomes.

In lieu of RCTs, adequate findings may be possible from long-term databases with prospectively collected data. However, these studies ideally should also use AS protocols that are

[g] See http://clinicaltrials.gov/ct2/show/NCT00499174; last accessed September 30, 2011.

defined *a priori* and undergo minimal change over time or between centers. A second-tier, but more practical option, is to allow for a range of *a priori* AS protocols across centers, also allowing for modifications over time; however, the AS protocols should be broadly similar to each other and any changes in protocols should be systematically tracked and adjusted for in analyses. The determination of which patients are potentially eligible for AS should also be made *a priori*. Only these patients (whether they ultimately received AS or another treatment) should be analyzed. To be interpretable, these studies will need to use multivariable analyses, propensity scores, or other validated methods (e.g., instrumental variable regression) to adjust for the broad range of factors that affect the decision to use AS. These include, but are not limited to disease factors (e.g., stage and grade); disease markers (e.g., PSA and imaging); patient demographics, psychometrics, personality traits, and personal relationships; clinic features and setting; primary care physician factors; and treating physician factors. We do not believe that retrospective studies (without *a priori* definitions of AS, eligibility criteria, or choice of variables of interest) are capable of having adequate data for unbiased analyses, because patient and tumor characteristics are strongly associated with initial treatment choice as well as outcomes (i.e., they are strong confounders of the treatment-outcome association).

Subgroup analyses of either the trials or the prospective comparative studies should be conducted to look for particular sets of men who may benefit most (or least) from one approach or the other. Preferably, these subgroups should be considered *a priori* to allow studies to be adequately powered for these subgroup analyses, to minimize bias, and to constrain type I error (false-positive findings). The factors listed in Key Questions 1 and 2 form a good starting point to consider which subgroups may be of interest. In addition, future studies that could uncover better bio- and imaging markers of indolent versus aggressive disease, and thus could better stage patients as having either low or high risk disease, are necessary to better inform which patients are most likely to benefit from observational versus active treatment.

Discussion

Prostate cancer epidemiology is affected by population-level trends, such as the aging of the United States population, but also by changes in the application of screening and diagnostic technologies among the population at risk. Keeping these caveats in mind, studies indicate that men in all racial/ethnic groups experienced increases in prostate cancer incidence since the mid-1980s. The incidence rate appears to have peaked in early-1990s. For all groups, incidence rates declined between the early-1990s and 1999. Studies consistently demonstrated that early-stage (localized and regional) prostate cancer cases were responsible for the observed increase in prostate cancer incidence from the mid-1980s up to the mid-1990s. Studies also demonstrated decreases in the prostate cancer-specific mortality rate for all age groups between the early-1990s and 1999. Mean age of diagnosis has also decreased over time from 72.2 years (1988 to 1989) to 67.2 years (2004 to 2005) for both blacks and whites. Another consistent trend over time in studies using the SEER database has been the decrease in low (corresponding to Gleason scores 2-4) and high grade (corresponding to Gleason score ≥7) tumors, and a concomitant increase in intermediate grade tumors (Gleason 5-6). It has been hypothesized that this effect is caused by changes in histopathological grading guidelines,[245] a preference towards avoiding assigning Gleason 2-4 scores based on prostate cancer biopsy samples[99,246,247], and the ability of the PSA test to detect moderately differentiated tumors with higher accuracy (compared to poorly-differentiated tumors). Most studies demonstrated decreasing trends in the proportion of patients being managed with strategies other than RP or RT throughout their respective time periods. Studies explicitly reporting on AS/WW-type strategies indicated decreases in the proportion of patients receiving such treatments over time; this was true even for subgroups of men with "low-risk disease."

There is not yet consistency among clinicians or researchers as to the definitions or standardizations of AS. Eligibility criteria for AS based on disease and patient characteristics and followup protocols including defining triggers for active interventions have not been standardized. This is apparent looking at the 16 unique cohorts with formal protocols for monitoring triggers for curative treatment of prostate cancer (AS cohorts). In all, a variety of observational management strategies was offered to men with low-risk or clinically localized prostate cancer although no uniform criteria were used to identify these men, with the exception that no cohorts enrolled patients with clinical stage greater than T2. The strategies included different combinations of periodic DRE, PSA testing, rebiopsy and/or imaging findings to determine different thresholds used for seeking definitive treatments. Additional information was provided by 13 unique cohorts of men who initially received no treatment and who were subsequently treated only for symptomatic progression (WW cohorts). About half of these WW cohorts were formed in the pre-PSA screening era, enrolled men with more advanced disease, and tended to use regular prostate acid phosphatase (PAP) testing in followup.

Because of the nonstandardized usages of the terms AS and WW coupled with the fact that the primary intents of the observational management strategies reviewed were frequently not reported, at times it was difficult when reviewing the studies to know who had AS and who had WW.

Only two studies specifically examined factors related to men who were enrolled in an active monitoring protocol with triggers for curative treatments. The first found that the free to total PSA ratio and T stage were independent predictors of time to radical treatments in patients on the protocol, while initial PSA, PSA density, Gleason score, number of positive cores, and prostate volume were not independent predictors. The second study found that men with decreased

baseline anxiety and higher socioeconomic status were associated with decreased probability of willingness to consent to randomization for AS versus definitive treatment (i.e., these men did not take a chance and proactively selected AS). The rest of the heterogeneous studies reported on men who did not receive treatments or initial treatments. Therefore, whether they were on AS or WW could not be readily discerned. The following patient and clinical variables are potentially important in increasing the probability that a patient receives WW or AS: older age, presence of comorbidities, lower Gleason score, lower tumor stage, lower diagnostic PSA, lower risk groups, or decreased baseline anxiety. The following patient and clinical variables are potentially important in increasing the probability that a patient interrupts WW or AS to seek definitive treatments: younger age, higher tumor stage, higher diagnostic PSA, higher PSA velocity, higher risk groups, or increased anxiety.

As most of these tentative conclusions are drawn from multivariable analyses of large databases that did not specifically address the factors that affect the offer, acceptance, and adherence of AS, whether different treatment options were offered to the patients, whether they accepted those options, and whether they adhered to their initial choices could only be inferred from whether they received the treatments or not. In addition, retrospective studies (without *a priori* definitions of AS, eligibility criteria, or choice of variables of interest) could not provide adequate data for unbiased analyses, because patient characteristics are strongly associated with initial treatment choice.

No trial has published results on comparisons of AS with RP, or RT in men with localized diseases. One trial reported that men on RP had lower mortality than men on WW; one trial reported that there was no difference in mortality comparing men in RP with men in WW. Retrospective studies suggest that men on conservative management had a higher prostate cancer-specific mortality than men treated with RP. Men who had RP had more urinary complications than men on WW. Retrospective studies also reported that men treated with RT had lower mortality than men on WW. They also reported higher rates of urinary strictures in men treated with RT compared with men on WW. It should be noted that confounding is likely in many retrospective analyses of large databases. The following example is instructive in illustrating the potential for confounding bias in observational studies of treatment effectiveness. Giordano 2008[248] replicated a previously published analysis[208] comparing overall survival between men who received active treatment (RP or RT) with men who were on WW, and performed additional analyses on mortality from non-prostate-specific causes (e.g., heart disease, other cancers, and chronic obstructive pulmonary disease). The study confirmed that patients with prostate cancer who underwent RP had better survival than those on WW but also suggested that patients treated with RP had improved survival compared to a randomly selected control population without cancer (matched with the RP population on the distributions of year of diagnosis and age at diagnosis). Further analysis showed that the patients who underwent RP had significantly improved their survival for many non-prostate-specific causes of death (including diabetes-, cardiovascular disease-, and chronic obstructive pulmonary disease-related death) as well, compared to those who had observational management. Because these causes of death are unlikely to be affected by prostate surgery, but conditions related to them (for example chronic angina or low exercise tolerance) are in fact used by urologists to decide fitness for surgery, the results suggest that the findings of improved overall survival in the RP group compared to the WW group reported by the previous analysis may have been affected by residual confounding by covariates that were not measured or not controlled for in the analysis.

Definitive conclusions for men with low-risk disease on AS or WW versus RP or RT will have to await results from two ongoing trials: Prostate cancer Intervention Versus Observation Trial (PIVOT: observation vs. RP[a]) and Prostate Testing for Cancer and Treatment trial (ProtecT: AS vs. RP or RT[b]). PIVOT recruited 731 patients (364 randomized to RP, 367 to observation); initial results were presented in the 2011 American Urological Association annual meeting (webcasts.prous.com/AUA2011/html/1-en/template.aspx?section=7&p=7,18082#). ProtecT has finished recruitment (>2500 men, about 700 men each were offered AS, RP, or RT) and is now in the followup phase.[c] A brief description of these studies is provided in Appendix Table B.

Although cost calculations using retrospective data were performed using different methods and followup durations in each study, overall it appears that WW is associated with lower treatment costs compared with active treatment. However, a cost analysis based on the ICER model indicates that with long-term followup, the costs of AS may exceed those of RP and BT; and may be lower than those of intensity modulated RT (IMRT) or proton beam RT.

Limitations in our approaches in this report largely concern the breadth of literature covered. For example, studies identified for the question on factors relevant to the practice of AS were primarily found from our literature search conducted specifically for AS, we did not do a general search on specific topics like insurance or patient compliance. Undoubtedly, there are studies on some of these factors in patients with other cancers and their findings could potentially be informative here.

In conclusion, more men are being diagnosed with early stage prostate cancer. Whether active monitoring with a curative intent is an appropriate option for these men remains unclear. A standard, universally agreed-upon definition of AS that clearly distinguishes it from WW and other observational management strategies is needed to help clarify scientific discourse in this field. Ongoing clinical trials may provide information on the comparative effectiveness of AS compared to immediate active treatment, but will require long term followup.

[a] Available at http://clinicaltrials.gov/ct2/show/study/NCT00007644; last accessed September 30, 2011.
[b] Available at http://clinicaltrials.gov/ct2/show/NCT00632983; last accessed September 30, 2011.
[c] See http://www.epi.bris.ac.uk/protect/pdffiles/newsletters/ProtecT%20Newsletter%20No8%20Dec%202010.pdf; last accessed September 30, 2011.

References

1. Siegel R, Ward E, Brawley O, et al. Cancer statistics, 2011: The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA Cancer J Clin 2011;61:212-236.
2. Ries LAG, Young JL, Keel, GE, et al. SEER Survival Monograph: Cancer Survival Among Adults: U.S. SEER Program, 1988-2001, Patient and Tumor Characteristics. Bethesda, MD: National Cancer Institute; 2007.
3. National Cancer Institute. Cancer Trends Progress Report - 2009/2010 Update. Bethesda, MD: National Institutes of Health; 2010.
4. Cooperberg MR, Lubeck DP, Mehta SS, et al. Time trends in clinical risk stratification for prostate cancer: implications for outcomes (data from CaPSURE). J Urol 2003;170:S21-S25.
5. Cooperberg MR, Grossfeld GD, Lubeck DP, et al. National practice patterns and time trends in androgen ablation for localized prostate cancer. J Natl Cancer Inst 2003;95:981-989.
6. Agency for Healthcare Research and Quality. Methods Reference Guide for Effectiveness and Comparative Effectiveness Reviews, Version 1.0. 2010. Ref Type: Generic
7. Martin RM, Gunnell D, Hamdy F, et al. Continuing controversy over monitoring men with localized prostate cancer: a systematic review of programs in the prostate specific antigen era. [Review] [45 refs]. Journal of Urology 2006;176:439-449.
8. Ip S, Dvorak T, Yu WW, et al. Comparative evaluation of radiation treatments for clinically localized prostate cancer: an update. Available at: http://www.cms.gov/mcd/viewtechassess.asp?where=index&tid=69. August 13, 2010. Ref Type: Generic
9. Wilt TJ, MacDonald R, Rutks I, et al. Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. Ann Intern Med 2008;148:435-448.
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264-9, W64.
11. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008-2012.
12. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
13. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62:1013-1020.
14. Altman DG, Bland JM. Presentation of numerical data. BMJ 1996;312:572.
15. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—agency for healthcare research and quality and the effective health-care program. J Clin Epidemiol 2010;63:513-523.
16. Mebane C, Gibbs T, Horm J. Current status of prostate cancer in North American black males. J Natl Med Assoc 1990;82:782-788.
17. Hankey BF, Feuer EJ, Clegg LX, et al. Cancer surveillance series: interpreting trends in prostate cancer—part I: Evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates. J Natl Cancer Inst 1999;91:1017-1024.
18. Chu KC, Tarone RE, Freeman HP. Trends in prostate cancer mortality among black men and white men in the United States. Cancer 2003;97:1507-1516.
19. Brawley OW. Prostate carcinoma incidence and patient mortality: the effects of screening and early detection. Cancer 1997;80:1857-1863.

20. Merrill RM, Brawley OW. Prostate cancer incidence and mortality rates among white and black men. Epidemiology 1997;8:126-131.

21. Farkas A, Schneider D, Perrotti M, et al. National trends in the epidemiology of prostate cancer, 1973 to 1994: evidence for the effectiveness of prostate-specific antigen screening. Urology 1998;52:444-448.

22. Perrotti M, Rabbani F, Farkas A, et al. Trends in poorly differentiated prostate cancer 1973 to 1994: observations from the Surveillance, Epidemiology and End Results database. J Urol 1998;160:811-815.

23. Dennis LK, Resnick MI. Analysis of recent trends in prostate cancer incidence and mortality. Prostate 2000;42:247-252.

24. Merrill RM, Lyon JL. Explaining the difference in prostate cancer mortality rates between white and black men in the United States. Urology 2000;55:730-735.

25. Clegg LX, Feuer EJ, Midthune DN, et al. Impact of reporting delay and reporting error on cancer incidence rates and trends. J Natl Cancer Inst 2002;94:1537-1545.

26. Stephenson RA. Prostate cancer trends in the era of prostate-specific antigen. An update of incidence, mortality, and clinical factors from the SEER database. Urol Clin North Am 2002;29:173-181.

27. Escobedo LG, Rivas SD, Holmes MD. Prostate cancer mortality in Connecticut, Iowa and New Mexico African American men. Cancer Detect Prev 2004;28:375-380.

28. McDavid K, Lee J, Fulton JP, et al. Prostate cancer incidence and mortality rates and trends in the United States and Canada. Public Health Rep 2004;119:174-186.

29. Hayat MJ, Howlader N, Reichman ME, et al. Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program. Oncologist 2007;12:20-37.

30. Jani AB, Johnstone PA, Liauw SL, et al. Age and grade trends in prostate cancer (1974-2003): a Surveillance, Epidemiology, and End Results Registry analysis. Am J Clin Oncol 2008;31:375-378.

31. Clegg LX, Li FP, Hankey BF, Chu K, Edwards BK. Cancer survival among US whites and minorities: a SEER (Surveillance, Epidemiology, and End Results) Program population-based study. Arch Intern Med. 2002;162:1985-1993.

32. Merrill RM, Morris MK. Prevalence-corrected prostate cancer incidence rates and trends. Am J Epidemiol 2002;155:148-152.

33. Collin SM, Martin RM, Metcalfe C, et al. Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study. Lancet Oncol 2008;9:445-452.

34. Sarma AV, Schottenfeld D. Prostate cancer incidence, mortality, and survival trends in the United States: 1981-2001. Semin Urol Oncol 2002;20:3-9.

35. Lu-Yao GL, Greenberg ER. Changes in prostate cancer incidence and treatment in USA. Lancet 1994;343:251-254.

36. Harlan L, Brawley O, Pommerenke F, et al. Geographic, age, and racial variation in the treatment of local/regional carcinoma of the prostate. J Clin Oncol 1995;13:93-100.

37. Jani AB, Master VA, Rossi PJ, et al. Grade migration in prostate cancer: an analysis using the Surveillance, Epidemiology, and End Results registry. Prostate Cancer Prostatic Dis 2007;10:347-351.

38. Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986-2005. J Natl Cancer Inst 2009;101:1325-1329.

39. Devesa SS, Blot WJ, Stone BJ, et al. A, Tarone RE, Fraumeni JF, Jr. Recent cancer trends in the United States. J Natl Cancer Inst. 1995;87:175-182.

40. Merrill RM, Stephenson RA. Trends in mortality rates in patients with prostate cancer during the era of prostate specific antigen screening. J Urol. 2000;163:503-510.

41. Miller DC, Gruber SB, Hollenbeck BK, et al. Incidence of initial local therapy among men with lower-risk prostate cancer in the United States. J Natl Cancer Inst 2006;98:1134-1141.

42. Shao YH, Demissie K, Shih W, et al. Contemporary risk profile of prostate cancer in the United States. J Natl Cancer Inst 2009;101:1280-1283.

43. Merrill RM, Potosky AL, Feuer EJ. Changing trends in U.S. prostate cancer incidence rates. J Natl Cancer Inst 1996;88:1683-1685.

44. Stewart SL, King JB, Thompson TD, et al. Cancer mortality surveillance--United States, 1990-2000. MMWR Surveill Summ 2004;53:1-108.

45. Zhu K, Devesa SS, Wu H, et al. Cancer incidence in the U.S. military population: comparison with rates from the SEER program. Cancer Epidemiol Biomarkers Prev 2009;18:1740-1745.

46. Polednak AP. Black-white differences in tumor grade (aggressiveness) at diagnosis of prostate cancer, 1992-1998. Ethn Dis 2002;12:536-540.

47. Underwood III W, Jackson J, Wei JT, et al. Racial treatment trends in localized/regional prostate carcinoma: 1992-1999. Cancer 2005;103:538-545.

48. Underwood W, De MS, Ubel P, et al. Racial/ethnic disparities in the treatment of localized/regional prostate cancer. J Urol 2004;171:1504-1507.

49. Demers RY, Swanson GM, Weiss LK, et al. Increasing incidence of cancer of the prostate. The experience of black and white men in the Detroit metropolitan area. Arch Intern Med 1994;154:1211-1216.

50. Severson RK, Montie JE, Porter AT, et al. Recent trends in incidence and treatment of prostate cancer among elderly men. J Natl Cancer Inst 1995;87:532-534.

51. Demers RY, Tiwari A, Wei J, et al. Trends in the utilization of androgen-deprivation therapy for patients with prostate carcinoma suggest an effect on mortality. Cancer 2001;92:2309-2317.

52. Schwartz KL, Grignon DJ, Sakr WA, et al. Prostate cancer histologic trends in the metropolitan Detroit area, 1982 to 1996. Urology 1999;53:769-774.

53. Gilliland FD, Hunt WC, Key CR. Ethnic variation in prostate cancer survival in New Mexico. Cancer Epidemiol Biomarkers Prev 1996;5:247-251.

54. Gilliland FD, Gleason DF, Hunt WC, et al. Trends in Gleason score for prostate cancer diagnosed between 1983 and 1993. J Urol 2001;165:846-850.

55. Newcomer LM, Stanford JL, Blumenstein BA, et al. Temporal trends in rates of prostate cancer: declining incidence of advanced stage disease, 1974 to 1994. J Urol 1997;158:1427-1430.

56. Stephenson RA, Smart CR, Mineau GP, et al. The fall in incidence of prostate carcinoma. On the down side of a prostate specific antigen induced peak in incidence—data from the Utah Cancer Registry. Cancer 1996;77:1342-1348.

57. Potosky AL, Miller BA, Albertsen PC, et al. The role of increasing detection in the rising incidence of prostate cancer. JAMA 1995;273:548-552.

58. Klabunde CN, Potosky AL, Harlan LC, et al. Trends and black/white differences in treatment for nonmetastatic prostate cancer. Med Care 1998;36:1337-1348.

59. Sheikh K, Bullock C. Rise and fall of radical prostatectomy rates from 1989 to 1996. Urology 2002;59:378-382.

60. Godley PA, Schenck AP, Amamoo MA, et al. Racial differences in mortality among Medicare recipients after treatment for localized prostate cancer. J Natl Cancer Inst 2003;95:1702-1710.

61. Zeliadt SB, Potosky AL, Etzioni R, et al. Racial disparity in primary and adjuvant treatment for nonmetastatic prostate cancer: SEER-Medicare trends 1991 to 1999. Urology 2004;64:1171-1176.

62. Carpenter WR, Howard DL, Taylor YJ, et al. Racial differences in PSA screening interval and stage at diagnosis. Cancer Causes Control 2010;21:1071-1080.

63. Mullins CD, Onukwugha E, Bikov K, et al. Health disparities in staging of SEER-medicare prostate cancer patients in the United States. Urology 2010;76:566-572.

64. Kindrick AV, Grossfeld GD, Stier DM, et al. Use of imaging tests for staging newly diagnosed prostate cancer: trends from the CaPSURE database. J Urol 1998;160:2102-2106.

65. Cooperberg MR, Lubeck DP, Grossfeld GD, et al. Contemporary trends in imaging test utilization for prostate cancer staging: data from the cancer of the prostate strategic urologic research endeavor. J Urol 2002;168:491-495.

66. Harlan SR, Cooperberg MR, Elkin EP, et al. Time trends and characteristics of men choosing watchful waiting for initial treatment of localized prostate cancer: results from CaPSURE. J Urol 2003;170:1804-1807.

67. Cooperberg MR, Lubeck DP, Meng MV, et al. The changing face of low-risk prostate cancer: trends in clinical presentation and primary management. J Clin Oncol 2004;22:2141-2149.

68. Cooperberg MR, Broering JM, Kantoff PW, et al. Contemporary trends in low risk prostate cancer: risk assessment and treatment. J Urol 2007;178:S14-S19.

69. Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate cancer. J Clin Oncol 2010;28:1117-1123.

70. Shah JB, McKiernan JM, Elkin EP, et al. Prostate biopsy patterns in the CaPSURE database: evolution with time and impact on outcome after prostatectomy. J Urol 2008;179:136-140.

71. Greene KL, Cowan JE, Cooperberg MR, et al. Who is the average patient presenting with prostate cancer? Urology 2005;66:76-82.

72. Mettlin CJ, Murphy G. The National Cancer Data Base report on prostate cancer. Cancer 1994;74:1640-1648.

73. Mettlin CJ, Murphy GP, Ho R, et al. The National Cancer Data Base report on longitudinal observations on prostate cancer. Cancer 1996;77:2162-2166.

74. Mettlin CJ, Murphy GP, McGinnis LS, et al. The National Cancer Data Base report on prostate cancer. American College of Surgeons Commission on Cancer and the American Cancer Society. Cancer 1995;76:1104-1112.

75. Mettlin CJ, Murphy GP, Rosenthal DS, et al. The National Cancer Data Base report on prostate carcinoma after the peak in incidence rates in the U.S. The American College of Surgeons Commission on Cancer and the American Cancer Society. Cancer 1998;83:1679-1684.

76. Danley KL, Richardson JL, Bernstein L, et al. Prostate cancer: trends in mortality and stage-specific incidence rates by racial/ethnic group in Los Angeles County, California (United States). Cancer Causes Control 1995;6:492-498.

77. Hamilton AS, Albertsen PC, Johnson TK, et al. Trends in the treatment of localized prostate cancer using supplemented cancer registry data. BJU Int 2011;107:576-584.

78. Smart CR. The results of prostate carcinoma screening in the U.S. as reflected in the surveillance, epidemiology, and end results program. Cancer 1997;80:1835-1844.

79. Merrill RM, Feuer EJ, Warren JL, et al. Role of transurethral resection of the prostate in population-based prostate cancer incidence rates. Am J Epidemiol 1999;150:848-860.

80. Hsing AW, Tsao L, Devesa SS. International trends and patterns of prostate cancer incidence and mortality. Int J Cancer 2000;85:60-67.

81. Mariotto AB, Wesley MN, Cronin KA, et al. Estimates of long-term survival for newly diagnosed cancer patients: a projection approach. Cancer 2006;106:2039-2050.

82. Abdollah F, Sun M, Thuret R, et al. A competing-risks analysis of survival after alternative treatment modalities for prostate cancer patients: 1988-2006. Eur Urol 2011;59:88-95.

83. Paltoo DN, Chu KC. Patterns in cancer incidence among American Indians/Alaska Natives, United States, 1992-1999. Public Health Rep 2004;119:443-451.

84. Kohler BA, Ward E, McCarthy BJ, et al. Annual report to the nation on the status of cancer, 1975-2007, featuring tumors of the brain and other nervous system. J Natl Cancer Inst 2011;103:714-736.

85. Schwartz KL, Severson RK, Gurney JG, et al. Trends in the stage specific incidence of prostate carcinoma in the Detroit metropolitan area, 1973-1994. Cancer 1996;78:1260-1266.

86. Barnholtz-Sloan JS, Severson RK, Vaishampayan U, et al. Survival analysis of distant prostate cancer by decade (1973-1997) in the Detroit Metropolitan Surveillance, Epidemiology and End Results (SEER) Program registry: has outcome improved? (United States). Cancer Causes Control 2003;14:681-685.

87. Warren JL, Yabroff KR, Meekins A, et al. Evaluation of trends in the cost of initial cancer treatment. J Natl Cancer Inst. 2008;100:888-897.

88. Gross CP, Smith BD, Wolf E, et al. Racial disparities in cancer therapy: did the gap narrow between 1992 and 2002? Cancer 2008;112:900-908.

89. Welch HG, Fisher ES, Gottlieb DJ, et al. Detection of prostate cancer via biopsy in the Medicare-SEER population during the PSA era. J Natl Cancer Inst 2007;99:1395-1400.

90. Cooperberg MR, Cowan J, Broering JM, et al. High-risk prostate cancer in the United States, 1990-2007. World J Urol 2008;26:211-218.

91. Mettlin CJ, Murphy GP, McDonald CJ, et al. The National Cancer Data base Report on increased use of brachytherapy for the treatment of patients with prostate carcinoma in the U.S. Cancer 1999;86:1877-1882.

92. White A, Coker AL, Du XL, et al. Racial/ethnic disparities in survival among men diagnosed with prostate cancer in Texas. Cancer 2011;117:1080-1088.

93. Frey CM, McMillen MM, Cowan CD, et al. Representativeness of the surveillance, epidemiology, and end results program data: recent trends in cancer mortality rates. J Natl Cancer Inst 1992;84:872-877.

94. Merrill RM, Dearden KA. How representative are the surveillance, epidemiology, and end results (SEER) program cancer data of the United States? Cancer Causes Control 2004;15:1027-1034.

95. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. Med Care 2002;40:IV-18.

96. Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst 2002;94:981-990.

97. Albertsen PC, Hanley JA, Barrows GH, et al. Prostate cancer and the Will Rogers phenomenon. J Natl Cancer Inst 2005;97:1248-1253.

98. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312:1604-1608.

99. Gofrit ON, Zorn KC, Steinberg GD, et al. The Will Rogers phenomenon in urological oncology. J Urol 2008;179:28-33.

100. Feuer EJ, Merrill RM, Hankey BF. Cancer surveillance series: interpreting trends in prostate cancer--part II: Cause of death misclassification and the recent rise and fall in prostate cancer mortality. J Natl Cancer Inst 1999;91:1025-1032.

101. Tsodikov A, Szabo A, Wegelin J. A population model of prostate cancer incidence. Stat Med 2006;25:2846-2866.

102. Etzioni R, Tsodikov A, Mariotto A, et al. Quantifying the role of PSA screening in the US prostate cancer mortality decline. Cancer Causes Control 2008;19:175-181.

103. Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. Biometrics 2008;64:10-19.

104. Draisma G, Postma R, Schroder FH, et al. Gleason score, age and screening: modeling dedifferentiation in prostate cancer. Int J Cancer 2006;119:2366-2371.

105. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst 2009;101:374-383.

106. Patel MI, DeConcini DT, Lopez-Corona E, et al. An analysis of men with clinically localized prostate cancer who deferred definitive therapy. J Urol 2004;171:1520-1524.

107. Al OM, Ross P, Fahmy N, et al. Role of repeated biopsy of the prostate in predicting disease progression in patients with prostate cancer on active surveillance. Cancer 2008;113:286-292.

108. Ercole B, Marietti SR, Fine J, et al. Outcomes following active surveillance of men with localized prostate cancer diagnosed in the prostate specific antigen era. J Urol 2008;180:1336-1339.

109. Eggener SE, Mueller A, Berglund RK, et al. A multi-institutional evaluation of active surveillance for low risk prostate cancer. J Urol 2009;181:1635-1641.

110. Soloway MS, Soloway CT, Williams S, et al. Active surveillance; a reasonable management alternative for patients with prostate cancer: the Miami experience. BJU Int 2008;101:165-169.

111. Dall'era MA, Konety BR, Cowan JE, et al. Active surveillance for the management of prostate cancer in a contemporary cohort. Cancer 2008;112:2664-2670.

112. Hardie C, Parker C, Norman A, et al. Early outcomes of active surveillance for localized prostate cancer. BJU Int 2005;95:956-960.

113. Ross AE, Loeb S, Landis P, et al. Prostate-specific antigen kinetics during follow-up are an unreliable trigger for intervention in a prostate cancer surveillance program. J Clin Oncol 2010;28:2810-2816.

114. Choo R, Deboer G, Klotz L, et al. PSA doubling time of prostate carcinoma managed with watchful observation alone. Int J Radiat Oncol Biol Phys 2001;50:615-620.

115. Adamy A, Yee DS, Matsushita K, et al. Role of prostate specific antigen and immediate confirmatory biopsy in predicting progression during active surveillance for low risk prostate cancer. J Urol 2011;185:477-482.

116. Metcalfe C, Tilling K, Davis M, et al. Current strategies for monitoring men with localised prostate cancer lack a strong evidence base: observational longitudinal study. Br J Cancer 2009;101:390-394.

117. San F I, Werner L, Regan MM, et al. Risk stratification and validation of prostate specific antigen density as independent predictor of progression in men with low risk prostate cancer during active surveillance. J Urol 2011;185:471-476.

118. Kakehi Y, Kamoto T, Shiraishi T, et al. Prospective evaluation of selection criteria for active surveillance in Japanese patients with stage T1cN0M0 prostate cancer. Jpn J Clin Oncol 2008;38:122-128.

119. Miocinovic R, Jones JS, Pujara AC, et al. Acceptance and durability of surveillance as a management choice in men with screen-detected, low-risk prostate cancer: improved outcomes with stringent enrollment criteria. Urology 2011;77:980-984.

120. van den Bergh RC, Vasarainen H, van der Poel HG, et al. Short-term outcomes of the prospective multicentre 'Prostate Cancer Research International: Active Surveillance' study. BJU Int 2010;105:956-962.

121. Newcomb LF, Brooks JD, Carroll PR, et al. Canary Prostate Active Surveillance Study: design of a multi-institutional active surveillance cohort and biorepository. Urology 2010;75:407-413.

122. Gorin MA, Soloway CT, Eldefrawy A, et al. Factors that influence patient enrollment in active surveillance for low-risk prostate cancer. Urology 2011;77:588-591.

123. Soloway MS, Soloway CT, Eldefrawy A, et al. Careful selection and close monitoring of low-risk prostate cancer patients on active surveillance minimizes the need for treatment. Eur Urol 2010;58:831-835.

124. Venkitaraman R, Norman A, Woode-Amissah R, et al. Prostate-specific antigen velocity in untreated, localized prostate cancer. BJU Int 2008;101:161-164.

125. Klotz L, Zhang L, Lam A, et al. A Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. J Clin Oncol 2010;28:126-131.

126. Fryback DG, Albertsen PC, Storer BE. Prostatectomy and survival among men with clinically localized prostate cancer. JAMA 1996;276:1723-1724.

127. Neulander EZ, Duncan RC, Tiguert R, et al. Deferred treatment of localized prostate cancer in the elderly: the impact of the age and stage at the time of diagnosis on the treatment decision. BJU Int 2000;85:699-704.

128. Venkitaraman R, Thomas K, Grace P, et al. Baseline urinary phytoestrogen levels and the natural history of untreated, localised prostate cancer in a British population. Int J Biol Markers 2008;23:192-197.

129. Krakowsky Y, Loblaw A, Klotz L. Prostate cancer death of men treated with initial active surveillance: clinical and biochemical characteristics. J Urol 2010;184:131-135.

130. Loblaw A, Zhang L, Lam A, et al. Comparing prostate specific antigen triggers for intervention in men with stable prostate cancer on active surveillance. J Urol 2010;184:1942-1946.

131. Cooperberg MR, Cowan JE, Hilton JF, et al. Outcomes of active surveillance for men with intermediate-risk prostate cancer. J Clin Oncol 2011;29:228-234.

132. Whitson JM, Porten SP, Hilton JF, et al. The relationship between prostate specific antigen change and biopsy progression in patients on active surveillance for prostate cancer. J Urol 2011;185:1656-1660.

133. Chan JM, Weinberg V, Magbanua MJ, et al. Nutritional supplements, COX-2 and IGF-1 expression in men on active surveillance for prostate cancer. Cancer Causes & Control 2011;22:141-150.

134. Giles SL, Morgan VA, Riches SF, et al. Apparent diffusion coefficient as a predictive biomarker of prostate cancer progression: value of fast and slow diffusion components. AJR 2011;American:586-591.

135. Porten SP, Whitson JM, Cowan JE, et al. Changes in prostate cancer grade on serial biopsy in men undergoing active surveillance. Journal of Clinical Oncology 2011;29:2795-2800.

136. Ng MK, Van AN, Thomas K, et al. Prostate-specific antigen (PSA) kinetics in untreated, localized prostate cancer: PSA velocity vs PSA doubling time. BJU Int 2009;103:872-876.

137. Carter HB, Kettermann A, Warlick C, et al. Expectant management of prostate cancer with curative intent: an update of the Johns Hopkins experience. J Urol. 2007;178:2359-2364.

138. Jones GW. Prospective, conservative management of localized prostate cancer. Cancer 1992;70:307-310.

139. Johansson JE. Expectant management of early stage prostatic cancer: Swedish experience. J Urol 1994;152:1753-1756.

140. Adolfsson J, Tribukait B, Levitt S. The 20-Yr outcome in patients with well- or moderately differentiated clinically localized prostate cancer diagnosed in the pre-PSA era: the prognostic value of tumour ploidy and comorbidity. Eur Urol 2007;52:1028-1035.

141. Handley R, Carr TW, Travis D, et al. Deferred treatment for prostate cancer. Br J Urol. 1988;62:249-253.

142. Rana A, Chisholm GD, Christodoulou S, et al. Audit and its impact in the management of early prostatic cancer. Br J Urol 1993;71:721-727.

143. Chen WM, Yang CR, Ou YC, et al. Clinical outcome of patients with stage T1a prostate cancer. J Chin Med Assoc 2003;66:236-240.

144. Holmberg L, Bill-Axelson A, Helgesen F, et al. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. N Engl J Med 2002;347:781-789.

145. Bangma CH, HOP WCJ, Schroder FH. Serial prostate specific antigen measurements and progression in untreated confined (stages T0 to 3NxM0, grades 1 to 3) carcinoma of the prostate. J Urol. 1995;154:1403-1406.

146. Wilt TJ, Brawer MK, Barry MJ, et al. The Prostate cancer Intervention Versus Observation Trial:VA/NCI/AHRQ Cooperative Studies Program #407 (PIVOT): design and baseline results of a randomized controlled trial comparing radical prostatectomy to watchful waiting for men with clinically localized prostate cancer. Contemp Clin Trials 2009;30:81-87.

147. Stratton MS, Reid ME, Schwartzberg G, et al. Selenium and inhibition of disease progression in men diagnosed with prostate carcinoma: study design and baseline characteristics of the 'Watchful Waiting' Study. Anticancer Drugs 2003;14:595-600.

148. McIntyre IG, Clarke RB, Anderson E, et al. . Molecular prediction of progression in patients with conservatively managed prostate cancer. Urology 2001;58:762-766.

149. Anai S, Nakamura K, Chang MN, et al. The feasibility of expectant management with inner-city men with newly diagnosed localized prostate cancer. J Health Care Poor Underserved 2008;19:164-170.

150. Roemeling S, Roobol MJ, de Vries SH, 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;51:1244-1250.

151. van den Bergh RC, Roemeling S, Roobol MJ, et al. Outcomes of men with screen-detected prostate cancer eligible for active surveillance who were managed expectantly. Eur Urol 2009;55:1-8.

152. Tilling K, Garmo H, Metcalfe C, et al. Development of a new method for monitoring prostate-specific antigen changes in men with localised prostate cancer: a comparison of observational cohorts. Eur Urol 2010;57:446-452.

153. Kakehi Y, Kamoto T, Ogawa O, et al. Clinical significance of nonpalpable prostate cancer with favorable biopsy features in Japanese men. Eur Urol 2000;37:552-558.

154. Arai Y, Egawa S, Kuwao S, et al. The role of volume-weighted mean nuclear volume in predicting tumour biology and clinical behaviour in patients with prostate cancer undergoing watchful waiting. BJU Int 2001;88:909-914.

155. Mohler JL, Williams BT, Freeman JA. Expectant management as an option for men with stage T1c prostate cancer: a preliminary study. World J Urol 1997;15:364-368.

156. Finelli A, Trottier G, Lawrentschuk N, et al. Impact of 5alpha-reductase inhibitors on men followed by active surveillance for prostate cancer. Eur Urol 2011;59:509-514.

157. McLaren DB, McKenzie M, Duncan G, et al. Watchful waiting or watchful progression?: Prostate specific antigen doubling times and clinical behavior in patients with early untreated prostate carcinoma. Cancer 1998;82:342-348.

158. Lee EK, Baack J, Penn H, et al. Active surveillance for prostate cancer in a veteran population. Can J Urol 2010;17:5429-5435.

159. Snyder CF, Frick KD, Blackford AL, et al. How does initial treatment choice affect short-term and long-term costs for clinically localized prostate cancer? Cancer 2010;116:5391-5399.

160. Dall'era MA, Hosang N, Konety B, et al. Sociodemographic predictors of prostate cancer risk category at diagnosis: unique patterns of significant and insignificant disease. J Urol 2009;181:1622-1627.

161. Barocas DA, Cowan JE, Smith Jr JA, et al. What percentage of patients with newly diagnosed carcinoma of the prostate are candidates for surveillance? An analysis of the CaPSURE database. J Urol 2008;180:1330-1334.

162. Konety BR, Cowan JE, Carroll PR. Patterns of primary and secondary therapy for prostate cancer in elderly men: analysis of data from CaPSURE. J Urol 2008;179:1797-1803.

163. Sadetsky N, Elkin EP, Latini DM, et al. Prostate cancer outcomes among older men: insurance status comparisons results from CaPSURE database. Prostate Cancer Prostatic Dis 2008;11:280-287.

164. Latini DM, Hart SL, Knight SJ, et al. The relationship between anxiety and time to treatment for patients with prostate cancer on surveillance. J Urol 2007;178:826-831.

165. Marr PL, Elkin EP, Arredondo SA, et al. Comorbidity and primary treatment for localized prostate cancer: data from CaPSURE. J Urol 2006;175:1326-1331.

166. Shavers VL, Brown ML, Potosky AL, et al. Race/ethnicity and the receipt of watchful waiting for the initial management of prostate cancer. J Gen Intern Med 2004;19:146-155.

167. Meng MV, Elkin EP, Harlan SR, et al. Predictors of treatment after initial surveillance in men with prostate cancer: results from CaPSURE. J Urol 2003;170:2279-2283.

168. Koppie TM, Grossfeld GD, Miller D, et al. Patterns of treatment of patients with prostate cancer initially managed with surveillance: results from The CaPSURE database. Cancer of the Prostate Strategic Urological Research Endeavor. J Urol 2000;164:81-88.

169. Moses KA, Paciorek AT, Penson DF, et al. Impact of ethnicity on primary treatment choice and mortality in men with prostate cancer: data from CaPSURE. J Clin Oncol 2010;28:1069-1074.

170. Latini DM, Elkin EP, Cooperberg MR, et al. Differences in clinical characteristics and disease-free survival for Latino, African American, and non-Latino white men with localized prostate cancer: data from CaPSURE. Cancer. 2006;106:789-795.

171. Wolters T, Roobol MJ, Steyerberg EW, et al. The effect of study arm on prostate cancer treatment in the large screening trial ERSPC. Int J Cancer 2010;126:2387-2393.

172. van As NJ, Norman AR, Thomas K, et al. Predicting the probability of deferred radical treatment for localised prostate cancer managed by active surveillance. Eur Urol 2008;54:1297-1305.

173. Mills N, Metcalfe C, Ronsmans C, et al. A comparison of socio-demographic and psychological factors between patients consenting to randomisation and those selecting treatment (the ProtecT study). Contemp Clin Trials 2006;27:413-419.

174. el-Geneidy M, Garzotto M, Panagiotou I, et al. Delayed therapy with curative intent in a contemporary prostate cancer watchful-waiting cohort. BJU Int 2004;93:510-515.

175. Wu H, Sun L, Moul JW, et al. Watchful waiting and factors predictive of secondary treatment of localized prostate cancer. J Urol 2004;171:1111-1116.

176. Carter CA, Donahue T, Sun L, et al. Temporarily deferred therapy (watchful waiting) for men younger than 70 years and with low-risk localized prostate cancer in the prostate-specific antigen era. J Clin Oncol 2003;21:4001-4008.

177. Yan Y, Carvalhal GF, Catalona WJ, et al. Primary treatment choices for men with clinically localized prostate carcinoma detected by screening. Cancer 2000;88:1122-1130.

178. Sommers BD, Beard CJ, D'Amico AV, Kaplan I, Richie JP, Zeckhauser RJ. Predictors of patient preferences and treatment choices for localized prostate cancer. Cancer 2008;113:2058-2067.

179. Meng MV, Elkin EP, Latini DM, et al. Treatment of patients with high risk localized prostate cancer: results from cancer of the prostate strategic urological research endeavor (CaPSURE). Journal of Urology 2005;173:1557-1561.

180. Roberts CB, Albertsen PC, Shao YH, et al. Patterns and correlates of prostate cancer treatment in older men. Am J Med 2011;124:235-243.

181. van den Bergh RC, Essink-Bot ML, Roobol MJ et al. Anxiety and distress during active surveillance for early prostate cancer. Cancer 2009;115:3868-3878.

182. Ramsey SD, Zeliadt SB, Fedorenko CR, et al. Patient preferences and urologist recommendations among local-stage prostate cancer patients who present for initial consultation and second opinions. World J Urol 2011;29:3-9.

183. Davison BJ, Oliffe JL, Pickles T, et al. Factors influencing men undertaking active surveillance for the management of low-risk prostate cancer. Oncol Nurs Forum 2009;36:89-96.

184. Durham J, Low M, McLeod D. Screening for prostate cancer: a survey of New Zealand general practitioners. N Z Med J 2003;116:U476.

185. Diefenbach MA, Dorsey J, Uzzo RG, et al. Decision-making strategies for patients with localized prostate cancer. Semin Urol Oncol 2002;20:55-62.

186. Zietman AL, Thakral H, Wilson L, et al. Conservative management of prostate cancer in the prostate specific antigen era: the incidence and time course of subsequent therapy. J Urol 2001;166:1702-1706.

187. Holmboe ES, Concato J. Treatment decisions for localized prostate cancer: asking men what's important. J Gen Intern Med 2000;15:694-701.

188. Fowler Jr FJ, McNaughton CM, Albertsen PC, et al. Comparison of recommendations by urologists and radiation oncologists for treatment of clinically localized prostate cancer. JAMA 2000;283:3217-3222.

189. Chapple A, Ziebland S, Herxheimer A, et al. Is 'watchful waiting' a real choice for men with prostate cancer? A qualitative study. BJU Int 2002;90:257-264.

190. Demark-Wahnefried W, Schildkraut JM, Iselin CE, et al. Treatment options, selection, and satisfaction among African American and white men with prostate carcinoma in North Carolina. Cancer 1998;83:320-330.

191. Berry DL, Ellis WJ, Woods NF, et al. Treatment decision-making by men with localized prostate cancer: the influence of personal factors. Urol Oncol 2003;21:93-100.

192. Steginga SK, Occhipinti S, Gardiner RA, et al. Making decisions about treatment for localized prostate cancer. BJU Int 2002;89:255-260.

193. Crawford ED, Bennett CL, Stone NN, et al. Comparison of perspectives on prostate cancer: analyses of survey data. Urology 1997;50:366-372.

194. O'Rourke ME. Narrowing the options: the process of deciding on prostate cancer treatment. Cancer Invest 1999;17:349-359.

195. Anandadas CN, Clarke NW, Davidson SE, et al. Early prostate cancer--which treatment do men prefer and why? BJU International 2011;107:1762-1768.

196. Xu J, Dailey RK, Eggly S, et al. Men's perspectives on selecting their prostate cancer treatment. Journal of the National Medical Association 2011;103:468-478.

197. Lin GA, Aaronson DS, Knight SJ, et al. Patient decision aids for prostate cancer treatment: a systematic review of the literature. CA Cancer J Clin 2009;59:379-390.

198. Jang TL, Bekelman JE, Liu Y, et al. Physician visits prior to treatment for clinically localized prostate cancer. Arch Intern Med 2010;170:440-450.

199. Shavers VL, Brown M, Klabunde CN, et al. Race/ethnicity and the intensity of medical monitoring under 'watchful waiting' for prostate cancer. Med Care 2004;42:239-250.

200. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA 1998;280:969-974.

201. Shariat SF, Karakiewicz PI, Roehrborn CG, et al. An updated catalog of prostate cancer predictive tools. Cancer 2008;113:3075-3099.

202. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1-12.

203. Auvinen A, Hakama M, la-Opas M, et al. A randomized trial of choice of treatment in prostate cancer: the effect of intervention on the treatment chosen. BJU Int 2004;93:52-56.

204. Davison BJ, Goldenberg SL, Wiens KP, et al. Comparing a generic and individualized information decision support intervention for men newly diagnosed with localized prostate cancer. Cancer Nurs 2007;30:E7-15.

205. Davison BJ, Degner LF. Empowerment of men newly diagnosed with prostate cancer. Cancer Nurs 1997;20:187-196.

206. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006;333:417.

207. Hadley J, Yabroff KR, Barrett MJ, et al. Comparative effectiveness of prostate cancer treatments: evaluating statistical adjustments for confounding in observational data. J Natl Cancer Inst 2010;102:1780-1793.

208. Wong YN, Mitra N, Hudes G, et al. Survival associated with treatment vs observation of localized prostate cancer in elderly men. JAMA 2006;296:2683-2693.

209. Albertsen PC, Hanley JA, Penson DF, et al. 13-year outcomes following treatment for clinically localized prostate cancer in a population based cohort. J Urol 2007;177:932-936.

210. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med 2005;352:1977-1984.

211. Iversen P, Madsen PO, Corle DK. Radical prostatectomy versus expectant treatment for early carcinoma of the prostate. Twenty-three year follow-up of a prospective randomized study. Scand J Urol Nephrol Suppl 1995;172:65-72.

212. Hegarty J, Beirne PV, Walsh E, et al. Radical prostatectomy versus watchful waiting for prostate cancer. Cochrane Database Syst Rev 2010;CD006590.

213. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med 2011;364:1708-1717.

214. Johansson E, Bill-Axelson A, Holmberg L, et al. Time, symptom burden, androgen deprivation, and self-assessed quality of life after radical prostatectomy or watchful waiting: the Randomized Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4) clinical trial. Eur Urol 2009;55:422-430.

215. Steineck G, Helgesen F, Adolfsson J, et al. Quality of life after radical prostatectomy or watchful waiting. N Engl J Med 2002;347:790-796.

216. Fransson P, Damber JE, Tomic R, et al. Quality of life and symptoms in a randomized trial of radiotherapy versus deferred treatment of localized prostate carcinoma. Cancer 2001;92:3111-3119.

217. Choo R, Long J, Gray R, et al. Prospective survey of sexual function among patients with clinically localized prostate cancer referred for definitive radiotherapy and the impact of radiotherapy on sexual function. Support Care Cancer 2010;18:715-722.

218. Elliott SP, Meng MV, Elkin EP, et al. Incidence of urethral stricture after primary treatment for prostate cancer: data From CaPSURE. J Urol 2007;178:529-534.

219. Zhou EH, Ellis RJ, Cherullo E, et al. Radiotherapy and survival in prostate cancer patients: a population-based study. Int J Radiat Oncol Biol Phys 2009;73:15-23.

220. Abdel-Wahab M, Reis IM, Hamilton K. Second primary cancer after radiotherapy for prostate cancer--a seer analysis of brachytherapy versus external beam radiotherapy. Int J Radiat Oncol Biol Phys 2008;72:58-68.

221. Tewari A, Divine G, Chang P, et al. Long-term survival in men with high grade prostate cancer: a comparison between conservative treatment, radiation therapy and radical prostatectomy—a propensity scoring approach. J Urol 2007;177:911-915.

222. Smith DP, King MT, Egger S, et al. Quality of life three years after diagnosis of localised prostate cancer: population based cohort study. BMJ 2009;339:b4817.

223. Stattin P, Holmberg E, Johansson JE, et al. Outcomes in localized prostate cancer: National Prostate Cancer Register of Sweden follow-up study. J Natl Cancer Inst 2010;102:950-958.

224. Thong MS, Mols F, Kil PJ, et al. Prostate cancer survivors who would be eligible for active surveillance but were either treated with radiotherapy or managed expectantly: comparisons on long-term quality of life and symptom burden. BJU Int 2010;105:652-658.

225. Fransson P, Damber JE, Widmark A. Health-related quality of life 10 years after external beam radiotherapy or watchful waiting in patients with localized prostate cancer. Scand J Urol Nephrol 2009;43:119-126.

226. Schapira MM, Lawrence WF, Katz DA, et al. Effect of treatment on quality of life among men with clinically localized prostate cancer. Med Care 2001;39:243-253.

227. Lubeck DP, Litwin MS, Henning JM, et al. Changes in health-related quality of life in the first year after treatment for prostate cancer: results from CaPSURE. Urology. 1999;53:180-186.

228. Bacon CG, Giovannucci E, Testa M, et al. The impact of cancer treatment on quality of life outcomes for patients with localized prostate cancer. J Urol 2001;166:1804-1810.

229. Litwin MS, Lubeck DP, Spitalny GM, et al. Mental health in men treated for early stage prostate carcinoma: a posttreatment, longitudinal quality of life analysis from the Cancer of the Prostate Strategic Urologic Research Endeavor. Cancer 2002;95:54-60.

230. Schymura MJ, Kahn AR, German RR, et al. Factors associated with initial treatment and survival for clinically localized prostate cancer: results from the CDC-NPCR Patterns of Care Study (PoC1). BMC Cancer 2010;10:152.

231. Berge V, Thompson T, Blackman D. Additional surgical intervention after radical prostatectomy, radiation therapy, androgen-deprivation therapy, or watchful waiting. Eur Urol 2007;52:1036-1043.

232. Rice KR, Colombo ML, Wingate J, et al. Low risk prostate cancer in men >/= 70 years old: To treat or not to treat. Urol Oncol 2011.

233. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990;1:43-46.

234. Greenland S, Robins JM. Empirical-Bayes adjustments for multiple comparisons are sometimes useful. Epidemiology 1991;2:244-251.

235. Penson DF, Schonfeld WH, Flanders SC, et al. Relationship of first-year costs of treating localized prostate cancer to initial choice of therapy and stage at diagnosis: results from the CAPSURE database. Urology 2001;57:499-503.

236. Wilson LS, Tesoro R, Elkin EP, et al. Cumulative cost pattern comparison of prostate cancer treatments. Cancer 2007;109:518-527.

237. Andersson SO, Andren O, Lyth J, et al. Managing localized prostate cancer by radical prostatectomy or watchful waiting: Cost analysis of a randomized trial (SPCG-4). Scand J Urol Nephrol 2011;45:177-183.

238. Ollendorf DA, Hayes JH, McMahon PM, et al. Management options for low-risk prostate cancer. 2009. Insitute for Clinical and Economic Review. Ref Type: Report

239. Ollendorf DA, Hayes JH, McMahon PM, et al. Active surveillance & radical prostatectomy for the management of low-risk, clinically-localized prostate cancer. January 11, 2009. Institute for Clinical and Economic Review. Ref Type: Report

240. Hayes JH, Ollendorf DA, Pearson SD, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. JAMA 2010;304:2373-2380.

241. Corcoran AT, Peele PB, Benoit RM. Cost comparison between watchful waiting with active surveillance and active treatment of clinically localized prostate cancer. Urology 2010;76:703-707.

242. Eldefrawy A, Katkoori D, Abramowitz M, et al. Active surveillance vs. treatment for low-risk prostate cancer: A cost comparison. Urol Oncol 2011.

243. Robins J, Orellana L, Rotnitzky A. Estimation and extrapolation of optimal treatment and testing strategies. Stat Med 2008;27:4678-4721.

244. Hernan MA, Lanoy E, Costagliola D, et al. Comparison of dynamic treatment regimes via inverse probability weighting. Basic Clin Pharmacol Toxicol 2006;98:237-242.

245. Srigley JR, Amin MB, Bostwick DG, et al. Updated protocol for the examination of specimens from patients with carcinomas of the prostate gland: a basis for checklists. Cancer Committee. Arch Pathol Lab Med 2000;124:1034-1039.

246. Thompson IM, Canby-Hagino E, Lucia MS. Stage migration and grade inflation in prostate cancer: Will Rogers meets Garrison Keillor. J Natl Cancer Inst 2005;97:1236-1237.

247. Epstein JI. Gleason score 2-4 adenocarcinoma of the prostate on needle biopsy: a diagnosis that should not be made. Am J Surg Pathol 2000;24:477-478.

248. Giordano SH, Kuo YF, Duan Z, et al. Limits of observational data in determining outcomes from cancer therapy. Cancer 2008;112:2456-2466.

249. Tseng KS, Landis P, Epstein JI, et al. Risk stratification of men choosing surveillance for low risk prostate cancer. J Urol 2010;183:1779-1785.

Abbreviations and Acronyms

3D-CRT	conformal RT
ADT	androgen deprivation therapy
AHRQ	Agency for Healthcare Research and Quality
AMSTAR	Assessment of Multiple Systematic Reviews tool
ANOVA	analysis of variance
AS	active surveillance
AUA	American Urological Association
BT	Brachytherapy
CaPSURE	Cancer of the Prostate Strategic Urologic Research Endeavor
CCRCT	Cochrane Central Register of Controlled Trials
CDC	Centers for Disease Control and Prevention
CDP	Consensus Development Program
CDSR	Cochrane Database of Systematic Reviews
CE	cost-effectiveness
CEA	cost-effectiveness analysis
DRE	digital rectal examination
EBRT	external beam RT
EPC	Evidence-based Practice Center
EPIC	Expanded Prostate Cancer Index Composite
Gy	Gray (unit of radiation dose)
HMO	health maintenance organization
HTA	Health Technology Assessment
ICER	Institute for Clinical and Economic Review
IMRT	intensity modulated RT
KQ	Key Question
NCDB	National Cancer Database
NCHS	National Center for Health Statistics
NCI	National Cancer Institute
NIH	National Institutes of Health
OS	overall survival
PAP	prostate acid phosphatase
PCOS	Prostate Cancer Outcomes Study
PIVOT	Prostate cancer Intervention Versus Observation Trial
POCS	Patterns of Care Study
PPO	preferred provider organization
ProtecT	Prostate Testing for Cancer and Treatment trial
PSA	prostate-specific antigen
QoL	quality of life
QoL-CS	quality of life-cancer survivors
RCT	randomized controlled trial
RP	radical prostatectomy
RT	radiation therapy (radiotherapy)
SEER	Surveillance, Epidemiology, and End Results Program of the NCI
TEP	Technical Expert Panel
TOO	Task Order Officer

TURP	transurethral resection of the prostate
TRUS	transrectal ultrasound
WW	watchful waiting

Appendix A. Search Strategies

A.1 Search Strategy for Systematic Reviews

Database: EBM Reviews - Cochrane Database of Systematic Reviews <2005 to December 2010>, EBM Reviews - Health Technology Assessment <4th Quarter 2010>, Ovid MEDLINE(R) <1996 to December Week 4 2010>
Search Strategy:

---

1 meta-analysis.pt.
2 systematic$ review$.mp.
3 (systematic$ adj9 overview$).mp.
4 (meta-analys$ or meta analys$ or metaanalys$).mp.
5 evidence review$.mp.
6 or/1-5
7 prostate cancer.mp.
8 watchful waiting.mp.
9 active surveillance.mp.
10 or/7-9
11 6 and 10

A.2 Search Strategy for Large Databases

Database: Ovid MEDLINE(R) (1996 to August Week 1 2011)

Search Strategy:

---

1. ("Surveillance, Epidemiology and End Results" or "SEER" or "National Cancer Data Base" or "NCDB" or "Cooperative Studies Program" or "CSP" or "CaPSURE").mp.
2. prostate cancer.mp. or exp Prostatic Neoplasms/
3. 1 and 2
4. ("Prostate Cancer Outcomes Study" or "PCOS").mp.
5. 2 and 4
6. 5 not 3

A.3 Search Strategy for AS or WW

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations (August 16, 2011), EBM Reviews - Cochrane Database of Systematic Reviews (2005 to July 2011), EBM Reviews - Cochrane Central Register of Controlled Trials (3rd Quarter 2011), EBM Reviews - Health Technology Assessment (3rd Quarter 2011), Ovid MEDLINE(R) without Revisions (1996 to August Week 1 2011)

Search Strategy:

---

1. watchful waiting.mp.
2. active surveillance.mp.
3. conservative management.mp.

4. expectant management.mp.
5. deferred treatment.mp.
6. ((expectant$ adj5 manage$) or (conservative$ adj5 manage$) or (active adj5 surveillance) or (watchful adj5 waiting) or (watch adj5 wait) or (watchful adj5 observation) or (active$ adj5 monitor$) or (defer$ adj5 treatment)).tw.
7. prostate cancer.mp. or exp Prostatic Neoplasms/
8. ((prostat$ adj5 neoplas$) or (prostat$ adj5 cancer$)).tw.
9. 1 or 2 or 3 or 4 or 5 or 6
10. 7 or 8
11. 9 and 10
12. limit 11 to English language
13. limit 12 to yr="1987-Current"
14. (expectant$ adj5 treatment).tw.
15. 10 and 14
16. limit 15 to yr="1987-Current"
17. 16 not 13

A.4 Search Strategy for the Cost-Effectiveness Analysis Registry

We identified all studies with the words “prostate” OR “prostatic” in the database. The search was performed on October 31, 2011. At that time the database contained records up to and including 2010.

Appendix B. Ongoing Randomized Studies Comparing Observational Management Strategies With Active Treatment Strategies for the Treatment of Clinically Localized Disease

Study name [Registration] Country	N centers (planned enrollment) [enrollment period]	Population	Intervention	Comparator(s)	Outcomes	Current status
PIVOT [Clinical Trials.gov, NCT00007644] USA	31 (731) [1994-2002]	Clinically localized prostate cancer, within 6 mo of diagnosis, ⩽75 yr	WW (expectant management with palliative therapy)	RP	OS (primary); prostate cancer-specific mortality, DFS, PFS, morbidity, QoL, and CE	Preliminary results presented at the 2011 AUA meeting
START [Clinical Trials.gov, NCT00499174] Canada, USA, UK	13 (2130) [2007-2011]	Histologically confirmed adenocarcinoma of the prostate, within 6 mo of diagnosis, clinical stage T1b-T2b, Gleason score ⩽ 6, PSA ⩽10 ng/mL, physical examination, DRE and transrectal US within 6 mo of randomization, radiographic studies, (if indicated) negative for metastasis, LE >10 yr	AS (PSA testing, repeat biopsy and DRE; radical intervention at biochemical, histological, or clinical progression)	RP or RT, based on patient and physician preference	Prostate cancer-specific mortality (primary); OS, QoL, distant DFS, PSA relapse/progression after radical intervention, initiation of ADT, proportion of patients on the AS arm who receive radical intervention , prognostic significance of PSA doubling-time prior to diagnosis, prognostic significance of molecular biomarkers	Terminated early (not meeting accrual target)
ProtecT [Clinical Trials.gov, NCT00632983] UK	10 (2050) [2001-ongoing]	clinically localized disease prostate cancer (T1-T2, NX, M0), 50-69 yr, PSA 3.0-19.99 ng/mL, no skeletal metastases by isotope bone scan, LE ⩾10 yr	AS (repeat PSA testing q3 mo in the first year and then q 6 mo thereafter; annual review appointment with DRE, if indicated)	RP and 3d-CRT (with or without ADT) (2 comparator arms)	OS (primary); disease progression, treatment complications, general health status, anxiety, depression, and psychological state, urinary symptoms, QoL, sexual function, qualitative evaluation of outcome by in-depth interviews	Followup phase

3d-CRT = 3-dimentional conformal radiotherapy; ADT = androgen deprivation therapy; AS = active surveillance; AUA = American Urological Association; CE = cost-effectiveness; DFS = disease-free survival; DRE = digital rectal examination; LE = life expectancy; mo = month; PIVOT = **P**rostate Cancer **I**ntervention **V**ersus **O**bservation **T**rial; NR = not reported; OS = overall survival; PFS, progression-free survival; ProtecT = **Pro**state **Te**sting for **C**ancer and **T**reatment; PSA = prostate specific antigen; QoL = quality of life; RP = radical prostatectomy; RT = radiation therapy; START = Active **S**urveillance **T**herapy **A**gainst **R**adical **T**reatment in Patients Diagnosed With Favorable Risk Prostate Cancer trial; US = ultrasound; yr = year.

Appendix C. Appendix Tables for Key Questions 1–4

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1[a]

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Mebane[16] 1990 2258952	SEER, TNCS	1969	1985	NR	Data presented only for Black and White individuals.
Hankey[17] 1999 10379964	SEER, NCHS	1969	1995	229,556 [prostate cancer cases in SEER, 1975-95]	Invasive prostate cancer cases, excluding histology codes for lymphoma and patients of unknown ethnicity. Data were available for 1968-95 for mortality and 1973-95 for incidence rates.
Chu[18] 2003 12627516	SEER, NCHS	1969	1999	NR	Men with a first primary prostate cancer. Data were available from 1969-99 for mortality and 1975-99 for incidence rates.
Brawley[19] 1997 9351560	SEER	1973	1994	NR	NR
Merrill[20] 1997 9229202	SEER	1973	1992	NR	NR
Smart[78] 1997 9351557	SEER	1973	1993	208,234	NR
Farkas[21] 1998 9730458	SEER	1973	1994	156,598	White and African-American prostate cancer patients
Perrotti[22] 1998 9720554	SEER	1973	1994	224,595	Excluded patients with tumors of anaplastic or transitional cell histology. Two groups of patients based on year of diagnosis (1980-84 and 1990-94) were assessed to reflect patterns of cancer presentation before and after the introduction of early detection methods (PSA and trans-rectal ultrasound guided biopsies).
Merrill[79] 1999 10522656	SEER	1973	1993	81,497 [cancers diagnosed between 1986 and 1993; data not reported for other years]	NR

[a] Pre-1980 data were not extracted for this review; however, in this table we have extracted information on the first and last year covered by each eligible study.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Dennis[23] 2000 10679753	SEER	1973	1996	253,833	Incident prostate cancer cases, aged ⩾45 yr. Cases ascertained from autopsy or death certificate only were excluded.
Hsing[b80] 2000 10585584	SEER	1973	1992	100,212 [90,419 white; 9,793 black]	NR
Merrill[24] 2000 10792091	SEER	1973	1995	NR	NR
Clegg[25] 2002 12381706	SEER	1973	1998	NR	Results were reported only for White and Black patients. The study reported age-adjusted and reporting delay-adjusted incidence rates for prostate cancer; reporting delay-adjusted rates were calculated only for 1981-98.
Stephenson[26] 2002 12109343	SEER	1973	1997	261,464	Incident prostate adenocarcinoma cases. Excluded cases not confirmed histologically and cases diagnosed by autopsy or death certificate only.
Escobedo[27] 2004 15542264	SEER	1973	1998	40,548	Patients in the SEER database who reside in Connecticut, Iowa, or New Mexico; African American men only.
McDavid[28] 2004 15192905	SEER	1973	1999	NR	Excluded men aged <50 yr and non-White individuals.
Hayat[29] 2007 17227898	SEER	1973	2003	134,434 [number available for the last study yr]	All cancer cases in the coverage areas (subgroup results presented for prostate cancer), excluding childhood cancers. Excluded were cancer diagnosed by death certificate only, or at autopsy, and cases with in situ disease as the first cancer diagnosis.
Jani[30] 2008 18845997	SEER	1974	2003	455,170	Patients with available grade and age information. Excluded stage IV disease, nonadenocarcinoma or undifferentiated histology, and those with missing data on grade or age.
Clegg[31] 2002 12230422	SEER	1975	1997	233,520	Incident invasive prostate cancer belonging to 6 ethnic groups: non-Hispanic Whites, Hispanic Whites, African Americans, Asian Americans (Chinese, Japanese, Filipino), American Indians and Hawaiian natives.
Merrill[32] 2002 11790678	SEER	1975	1997	NR	Black or White prostate cancer patients.

[b] Incidence rates and patient characteristics in this study were calculated from SEER data and were extracted; mortality rates were obtained from published sources and were not extracted.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Mariotto[81] 2006 16572414	SEER	1975	2000 [cases diagnosed up to 1999]	NR	The study also reported data and projections based on the Connecticut tumor registry. These analyses were not stratified by factors relevant to Key Question 1 of the report and were not extracted in these evidence tables.
Collin[33] 2008 18424233	SEER	1975	2004	NR	NR
Sarma[34] 2002 11828352	SEER	1981	1998	NR	NR
Lu-Yao[35] 1994 7905093	SEER	1983	1989	NR	Whites prostate cancer patients, 50-79 yr
Harlan[36] 1995 7799048	SEER	1984	1991	67,693	Localized or regional prostate cancer; excluded cases identified by death certificate only or at autopsy.
Jani[37] 2007 17505529	SEER	1984	2003	411,325	Excluded patients with stage IV disease, non-adenocarcinoma histology, missing information on stage or grade, undifferentiated disease.
Welch[38] 2009 19720969	SEER	1986	2005	NR	NR
Devesa[39] 1995 7707404	SEER	1987	1991	NR	Excluded in situ cancers. Rates during the period 1975-1979 were used as baseline.
Merrill[40] 2000 10647666	SEER	1988	1995	64,562 [analyses by age and stage at diagnosis were based on 64,455 and 64,463 individuals, respectively]	Men diagnosed with prostate cancer between 1973 and 1995; data reported on causes of death during the period 1988-95. Analyses stratified by age excluded patients <50 yr and analyses stratified by stage excluded patients with in situ tumors.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Miller[41] 2006 16912266	SEER	1988	2002	71,602 [an additional historical cohort of 25,826 men diagnosed during 1988-90 was used to compare treatment patterns]	Localized or regional adenocarcinoma of the prostate. Excluded men with missing information that precluded assigned to a risk group (based on age and grade criteria) and those with missing information on primary treatment, race/ethnicity, or marital status.
Shao[42] 2009 19713548	SEER	1988	2005	82,541 (2004-05) (NR for other yr)	Age ≥25 yr. Excluded men of race/ethnicity other than Black or White, those with missing data on age, PSA, Gleason score, or clinical stage
Abdollah[82] 2011 20965646	SEER	1988	2006	404,604	Histologically-confirmed nonmetastatic prostatic adenocarcinoma, managed with RT, RP or observation. Only included men with localized prostate cancer, aged 30-95 yrs. Excluded individuals with unknown grade, or undifferentiated disease.
Merrill[43] 1996 8931614	SEER	1989	1993	80,936 [72,659 White; 8277 Black]	White and Black prostate cancer patients.
Stewart[44] 2004 15179359	SEER, NCHS	1990	2000	NR	NR
Zhu[45] 2009 19505907	SEER	1990	2004	42,751	NR [The study reported additional data from ACTUR, a tumor registry for military personnel; these data were not considered representative of the US population and pertained to <1000 prostate cancer cases, thus they were not extracted.]
Polednak[46] 2002 12477140	SEER	1992	1998	46,248 [in 1992: 2969 Black, 23,347 non-Hispanic White; in 1997: 2821 Black, 17,111 non-Hispanic White	Excluding cases ascertained from death certificate only or autopsy. Data only reported for non-Hispanic White and Black patients.
Paltoo[83] 2004 15219802	SEER	1992	1999	NR	Any incident cancer. American Indian or Alaskan Native, or white.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Underwood[48] 2004 15017208	SEER	1992	1999	142,340	Localized or regional, histologically-confirmed adenocarcinoma. Patients with missing data on race, treatment received or tumor grade were excluded. Analyses were restricted to the 3 largest race/ethnicity groups in SEER (White, Hispanic, African American).
Underwood[47] 2005 15612083	SEER	1992	1999	142,340	Localized or regional, histologically-confirmed adenocarcinoma. Patients with missing data on race, treatment received or tumor grade were excluded. Analyses were restricted to the 3 largest race/ethnicity groups in SEER (White, Hispanic, African American).
Kohler[84] 2011 21454908	SEER, NPCR, NAACCR	1998	2007	NR	This is the most recent "Annual Report to the Nation on the Status of Cancer"; the report presents data on incident invasive cancer cases from population-based cancer registries that are NACCR members and participate in the NCI's SEER program or the CDC's NPCR.
Demers[49] 1994 8203988	SEER, Detroit	1973	1991	22,632	Incident prostate cancer cases; only included White and Black individuals.
Severson[50] 1995 7707440	SEER, Detroit	1973	1992	12,413	Prostate cancer patients aged ⩾75 yr
Schwartz[85] 1996 8826949	SEER, Detroit	1973	1994	44,302	Metropolitan Detroit Cancer Surveillance System only. Incident prostate cancer, black or white men.
Demers[51] 2001 11745285	SEER, Detroit	1973	1998	56,425	Residents of Wayne, Oakland, or Macomb Counties who died due to prostate cancer
Barnholtz-Sloan[86] 2003 14575366	SEER, Detroit	1973	2001 [cases diagnosed up to 1997]	24,342	Patients with newly diagnosed primary distant stage prostate cancer, registered in the Metropolitan Detroit SEER registry; Caucasian or African American.
Schwartz[52] 1999 10197854	SEER, Detroit	1982	1996	39,566	NR
Gilliland[53] 1996 8722215	SEER, New Mexico	1983	1992	7563	Histologically-confirmed incident prostate cancer
Gilliland[54] 2001 11176484	SEER, New Mexico	1983	1993	1535	Histologically-confirmed prostate adenocarcinoma

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Newcomer[55] 1997 9302136	SEER, Seattle-Puget Sound	1974	1994	33,086	White or African-American men, aged ⩾35 yr with incident histologically-confirmed prostate cancer. Excluded cases identified through autopsy report only.
Stephenson[56] 1996 8608513	SEER, Utah	1984	1993	8867 (different analytic samples were used for different analyses)	NR
Potosky[57] 1995 7530782	SEER-Medicare	1986	1991	NR	Prostate cancer patients aged ⩾65 yr.
Klabunde[58] 1998 9749657	SEER-Medicare	1986	1993	52,915	Black and White men, aged 65 or older, diagnosed with prostate cancer. Excluded men whose tumor stage was recorded as in situ, distant, or unstaged, those with unknown tumor grade, those diagnosed with prostate cancer by autopsy or only on death certificates, those with incomplete Medicare claims data and those with racial/ethnic classification other than White or Black.
Sheikh[59] 2002 11880074	SEER-Medicare	1986	1996	NR	Medicare beneficiaries aged ⩾65 yr (for population rates); incident cases of prostate cancer (to define the population at risk of undergoing radical prostatectomy)
Godley[60] 2003 14625261	SEER-Medicare	1986	1996	43,989	Localized prostate cancer, aged 65-84 yr, enrolled in Medicare Part A or Part B for at least one month of the study period. Patients were excluded if they had a prior cancer diagnosis, a second cancer diagnosed in the same month as prostate cancer, had non-invasive disease, were missing the month of diagnosis, were aged <65 yr, were neither Black or nor non-Hispanic White, were diagnosed at death, had no Medicare coverage during the study period, or were aged ⩾85 yr at diagnosis. Patients with locally advanced, metastatic or unstaged cancer were also excluded.
Zeliadt[61] 2004 15596192	SEER-Medicare	1991	1999	90,128	White or African-American men, aged ⩾65 with incident prostate cancer, clinically staged as localized or regional, eligible for Medicare Part A and Part B, not enrolled in HMO at diagnosis. Patients with metastatic disease, those who became ineligible for Medicare during the followup period and those with an orchiectomy claim more than 3 months before diagnosis were excluded.
Warren[87] 2008 18544740	SEER-Medicare	1991	2002	106,534 [patients with claims for specific cancer services]	Men aged ⩾65 diagnosed with prostate cancer stage I-IV, enrolled in Medicare Part A and Part B coverage for 2 months before and 12 months after their diagnosis (i.e. Medicare claims up to 2003 were considered). Excluded patients enrolled in managed care at any time during this 14-month period.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Gross[88] 2008 18181101	SEER-Medicare	1992	2002	82,328 (74,288 white, 8040 black)	Patients with early stage prostate cancer who received RP, BT, or EBRT; blacks and whites only. Data on treatment patterns were reported only among patients with localized disease.
Welch[89] 2007 17848671	SEER-Medicare	1993	2001	10,429 [men with a needle biopsy of the prostate]	Men with a claim for a needle biopsy of the prostate during the study period were included, base on the 5% Medicare Part B sample. Men not aged ⩾65 yr at the start of each calendar year, not entitled to Part B for the entire year, or enrolled in risk-contract managed care plans were excluded. Men who had a concomitant transurethral resection of the prostate or who had a prior prostate cancer diagnosis were also excluded.
Carpenter[62] 2010 20333462	SEER-Medicare	1994	2002	18,067	Black and White men, aged ⩾65 yr, diagnosed with prostate cancer in the following 8 SEER registries: Atlanta, Connecticut, Detroit, Rural Georgia, Los Angeles, San Francisco-Oakland, San Jose, Seattle-Puget Sound. Included patients had to have known prostate cancer stage, no prior or concurrent cancers, no gaps in Medicare coverage for 3 yr prior to diagnosis, and no HMO coverage from enrollment to diagnosis.
Mullins[63] 2010 20163844	SEER-Medicare	1998	2002	42,318 [13,447 in 1998; 28,871 in 2002]	Prostate cancer patients with Medicare, aged ⩾65 yr. Excluded men with unknown mo of diagnosis, <65 yr old at time of diagnosis, races other than non-Hispanic White, AA, and White Hispanic, missing Census Tract information.
Kindrick[64] 1998 9817332	CaPSURE	1989	1997	3557	Biopsy-confirmed prostate adenocarcinoma. Excluded patients with a missing date of diagnosis or those for whom no initial treatment was recorded. Patients included in CaPSURE between 1995-97 were included, including non-incident cases (those diagnosed since 1989).
Cooperberg[65] 2002 12131295	CaPSURE	1989	2001	4966	Excluded those diagnosed before 1989 and those with missing information on primary treatment or clinical staging.
Cooperberg[4] 2003 14610406	CaPSURE	1989	2002	6290	Unselected men with biopsy-proven prostate adenocarcinoma. Excluded those with missing data on PSA, T stage or multiple parameters.
Cooperberg[5] 2003 12837834	CaPSURE	1989	2001	3439	Patients who received RP, EBRT, BT, PADT or WW as primary therapy. Patients with incomplete clinical staging information, those with missing data on treatment and those receiving cryotherapy as primary treatment were excluded.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Harlan[66] 2003 14532780	CaPSURE	1989	2000	5365 (402 received WW)	Patients with biopsy-confirmed prostate adenocarcinoma, localized stage (T3a or lower, N0 M0) who chose WW or active treatment within 9 mo of diagnosis. Patients who waited more than 9 mo after diagnosis before initiating active treatment and those who received active treatment before or within 6 mo after initiating WW were excluded from the analysis. Active treatment was defined as RP, EBRT, interstitial RT, cryotherapy, or ADT.
Cooperberg[67] 2004 15169800	CaPSURE	1989	2001	1990	Excluded patients with unknown PSA at diagnosis, diagnostic biopsy Gleason score, and/or clinical T stage. Excluded patients with unknown primary treatment, and those receiving cryotherapy as primary therapy (this group accounted for 2% of patients since 1996; 68% of cryotherapy treated patients were treated at a single practice site in the early 1990s). Only low-risk prostate cancer patients were analyzed for the temporal trends of clinical characteristics or treatments received.
Cooperberg [68] 2007 17644125	CaPSURE	1990	2006	10,385	Excluded those diagnosed before 1990, those with metastatic or locally advanced disease (clinical stage T3b or higher) and those with missing data on PSA, T stage or biopsy Gleason score. Localized biopsy-proven prostate adenocarcinoma. Information relevant to this review is only reported for "low-risk" prostate cancer, defined as PSA$\leqslant$10 ng/ml, Gleason score$\leqslant$6 and clinical stage T1/2a.
Cooperberg[90] 2008 18369637	CaPSURE	1990	2007	10,808 [3372 were considered high risk]	Localized prostate cancer, excluding metastatic or locally advanced ($\geqslant$T3bN0M0) disease, missing PSA, clinical T stage, or Gleason score data. Results for trends over time were reported only among the high risk subset of patients. High risk was defined as stage T2c/T3a, PSA >20ng/mL, or Gleason score $\geqslant$8.
Cooperberg[69] 2010 20124165	CaPSURE	1990	2008	11,892	Biopsy-proven prostate cancer. Excluded advanced disease (stage higher than T3a N0 M0); diagnosed before 1990; those from sites contributing <30 pts; and those receiving treatments other than RP, EBRT, BT, cryoablation, WW/AS or ADT or with unknown primary treatment.
Shah[70] 2008 17997437	CaPSURE	1995	2004	6450	Unselected men with biopsy-proven prostate adenocarcinoma. Enrolled in CaPSURE within 6 mo of diagnosis with complete clinical information (PSA, Gleason score and clinical stage) and complete followup information. Excluded patients with fewer than 6 biopsy cores or unavailable biopsy details.
Greene[71] 2005 16194711	CaPSURE	1997	2003	3003	Patients with biopsy-proven prostate cancer; availability of pretreatment demographics and QoL data.

Appendix Table C1.1. Descriptive characteristics of the epidemiologic studies considered relevant to KQ1 (continued)

Author Year uid	Database	Start year	End year	Sample size	Eligibility criteria
Mettlin[72] 1994 8062197	NCDB	1985	1990	85,813	NR
Mettlin[73] 1996 8640686	NCDB	1985	1993	349,154	Convenience sample of cancer patients in hospitals that voluntarily participated in the database.
Mettlin[74] 1995 8625214	NCDB	1986	1992	108,717 [number of patients included in analyses of stage by time period; the analytic sample size varied between analyses]	All submitted data to NCDB in 1986-1987 and in 1992. The available data represented ~30% and 77% of all prostate cancers diagnosed in the US in 1986/1987 and 1992, respectively.
Mettlin[75] 1998 9781963	NCDB	1992	1995	176,316	NR
Mettlin[91] 1999 10547563	NCDB	1992	1996	435,264	Prostate carcinoma (as reported from participating hospitals).
Danley[76] 1995 8580296	LAC/USC CSP	1976	1988	29,992	Invasive prostate cancer patients, aged ⩾45 yr. Excluded men younger than 45 yr of age.
Hamilton[77] 2011 20735387	POCS, NCI	1998	2002	2101 [962 in 1998 and 1139 in 2002; the weighted sample size was 15,547 for 1998 and 31,367 in 2002]	The 1998 POCS included primary histologically confirmed prostate cancer patients diagnosed between Jan 1st, 1998 and Dec 31st, 1998. Patients with a history of previous cancer (except non-melanoma skin cancer) and those aged <21 yr were excluded. A stratified random sample of all prostate cancer cases diagnosed in 10 regional population based cancer registries were selected, according to patients' race/ethnicity and age at diagnosis. For the 2002 POCS, similar procedures were followed as in 1998 with the inclusion of additional registries.
White[92] 2011 21351084	Texas Cancer Registry	1995	2002	87,444 [87,449 in abstract]	Incident invasive prostate cancer cases in men ⩾25 yrs (last date of follow-up was December 31, 2003)

Studies are arranged by database, then chronologically by the first year of enrollment, then by year of publication.
ACTUR = Automated Central Tumor Registry; ADT = androgen deprivation therapy; BT = brachytherapy; EBRT = external beam radiation therapy; HMO = health maintenance organization; mo = months; NCHS = National Center for health Statistics; NR = not reported; POCS = Patterns of Care; PSA = prostate-specific antigen; QoL = quality of life; RT = radiation therapy; RP = radical prostatectomy; TNCS = Third National Cancer Survey; VA = Veteran's Administration; WW = watchful waiting; yr = year.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Patient characteristics – Age										
Hankey[17] 1999 10379964	SEER	50-59 60-69 70-79 ≥80 *Annual percentage change of the age-standardized (1970 US standard) incidence rates*	NA 2.3 (1.5, 3.1) (1973-84) NA NA	3.8 (2.7, 4.9) (1973-89) 8.4 (6.1, 10.8) (1984-89) 3.3 (2.5, 4.1) (1973-88) 0.8 (-0.2, 1.9) (1973-85)	32.7 (12.5, 56.4) (1989-92) 24.2 (17.8, 30.9) (1989-92) 16.5 (10.7, 22.7) (1988-92) 6.4 (3.6, 9.3) (1985-92)	1.9 (-4.2, 8.5) (1992-95) -9.0 (-11.2, -6.7) (1992-95) -16.9 (-21.1, -12.6) (1992-95) -22.1 (-27.1, -16.8) (1992-95)			Annual percentage changes calculated from joinpoint regression (95% CI).	Additional information is presented in Figure 2 of the paper.
Merrill[79] 1999 10522656	SEER	65-69 70-74 75-79 ≥80 ≥65 *Annual change in age-adjusted incidence rates (1970 world standard)*	(1973)	(1986)					+3.25 (2.63, 3.87) +2.83 (2.43, 3.23) +2.12 (1.78, 2.46) +1.32 (0.90, 1.75) +2.27 (1.99, 2.56) [Annual percentage change and corresponding 95% CI]	

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Escobedo[27] 2004 15542264	SEER-Connecticut	<54 55-54 65-74 ≥75 *Age-adjusted (1970 US standard) rates*	3.8 210.3 721.5 1157.3 (1973-88)		13.7 567.2 1450.0 1495.2 (1989-98)				NR	African American men only
	SEER-Iowa	<54 55-54 65-74 ≥75 *Age-adjusted (1970 US standard) rates*	4.0 211.7 654.2 1275.1 (1973-88)		13.0 512.1 1253.6 1589.9 (1989-98)				NR	African American men only
	SEER-New Mexico	<54 55-54 65-74 ≥75 *Age-adjusted (1970 US standard) rates*	3.5 257.7 475.5 1195.4 (1973-88)		6.7 295.0 1093.7 978.3 (1989-98)				NR	African American men only
McDavid[28] 2004 15192905	SEER	50-59 60-69 70-79 ≥80 *Annual percentage change for each time period.*	NA 3.3* (1973-85) 2.6* (1973-86) 1.5* (1973-87)	4.2* (1973-89) 9.6* (1985-89) 12.5* (1986-92) 7.0* (1987-92)	29.6* (1989-92) 23.88 (1989-92) -14.5* (1992-95) -19.8* (1992-95)	4.1* (1992-99) -7.9* (1992-95) 0.6 (1995-99) -1.8 (1995-99)	NA 1.5* (1995-99) NA NA		* Denotes P<0.05. P-values were not provided for the other estimates.	The reporting of specific time intervals was determined by joint-point analysis. Additional information is presented in Figure 3 of the paper.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Lu-Yao[35] 1994 7905093	SEER	50-59 60-69 70-79 *Annual percent change in incidence*		+5.2%/yr +7.7%/yr +5.7%/yr (1983-89)					NR	
Welch[38] 2009 19720969	SEER	<50 50-59 60-69 70-79 ≥80 *Incidence rate*		1.3 58.4 349.4 819.2 1146.5 (1986)				9.4 212.7 666.9 896.8 637.4 (2005)	7.23 (6.4-8.2) 3.64 (3.3-4.0) 1.91 (1.8-2.0) 1.09 (1.05-1.14) 0.56 (0.53-0.60) [RR (95% CI), comparing 1986 to 2005]	Additional information (1986-2005) is presented in Figure 1 of the paper.
Devesa[39] 1995 7707404	SEER	35-54 55-74 ≥75 *Age-adjusted (1970 US standard) incidence rate*			13.2 459.7 1278.1 (1987-91)				+6.0 (+83.3%) +215.7 (+88.4%) +365.3 (+40%) Absolute [relative] change in rate, compared to 1975-79	
Merrill[43] 1996 8931614	SEER	50-59 60-69 70-79 80+ *Percent change in age-adjusted (1990 US standard) prostate cancer incidence*			+4% -9% -20% -29% (1992-93)				NR	White males only (n=72,659)

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	50-59 60-69 70-79 80+ *Percent change in age-adjusted (1990 US standard) rate*			+25% +15% -8% -17% (1992-93)				NR	Black males only (n=8,277)
Newcomer[55] 1997 9302136	SEER, Seattle-Puget Sound	35-64 ≥65 Age-adjusted (1970 US standard) rates	42 NR (1974)		159 NR (1991)				+279% +163% (change in incidence, 1974 vs. 1991)	Additional information is presented in Figure 4 of the paper.
Potosky[57] 1995 7530782	SEER-Medicare	65-74 ≥75		685 1141 (1987)	1175 1615 (1991)				Number of new cases per 100,000 men, age-adjusted using the direct method to the 1970 US standard population	Men ≥65 yr old
Sheikh[59] 2002 11880074	SEER-Medicare	65-74 ≥75 *Incidence rate*		746 1212 (1989)	1323 1747 (1992)	899 920 (1996)			NR	Additional data are presented in Figure 3 of the paper (1986-96)

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Patient characteristics – Race/ethnicity										
Mebane[16] 1990 2258952	SEER	White Black *Age-adjusted (1970 US standard) incidence rates*	68.9 126.0 (1980)	75.9 127.4 (1985)					NR	
Hankey[17] 1999 10379964	SEER	Whites Black Other *Annual percentage change of the age-standardized (1970 US standard) incidence rates*	2.4 (1.8, 3.1) (1973-85) NA NA	6.9 (2.5, 11.6) (1985-89) 2.3 (1.8, 2.8) (1973-89) 1.5 (.3, 2.7) (1973-1987)	18.4 (10.7, 26.6) (1989-92) 17.0 (12.4, 21.8) (1989-93) 16.2 (10.8, 21.9) (1987-92)	-12.8 (−15.7, −9.8) (1992-95) -14.0 (−20.2, −7.4) (1993-95) -7.5 (-12.9, -1.7) (1992-95)			Annual percentage changes calculated from joinpoint regression (95% CI).	Additional information is presented in Figure 1 of the paper.
Brawley[19] 1997 9351560	SEER	White Black *Age-adjusted incidence rate (standardized to 1970 values)*	(1973)		(1994)				Change in incidence = +130% among White Americans, peak in 1992; +140% among Black Americans, peak in 1993	

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Farkas[21] 1998 9730458	SEER	Whites			Peak, +0.43 (1991) Start of decline, −0.28 (1991)				NR	
		African-Americans *[Rate of increase in organ confined disease incidence per 100,000 persons]*			Peak, +0.73 (1992) Start of decline, −0.29 (1994)					
Hsing[80] 10585584 2000	SEER	White Black *Age-adjusted (world standard) rates*	47.9 79.9 (1973-77)		100.8 137.0 (1988-92)				+110.4% +71.5% (% change between periods)	No data were reported for other ethnicities.
Clegg[25] 2002 12381706	SEER	White *Estimated annual percentage change in age-adjusted (1970 US standard) incidence rates*		-3.2 (2.7, 3.7) (1973-87)	20.7 (14.5-27.3) (1988-92) -13.5 (−23.4-, −2.3) (1993-94)	2.2 (-2.8-7.4) (1995-98)			Estimates from joinpoint regression (95% CI)	The study did not report estimates for other racial/ethnic groups.
	SEER	White *Estimated annual percentage change in age-adjusted (1970 US standard) reporting delay-adjusted incidence rates*		3.2 (2.7, 3.8) (1973-87)	19.9 (14.3, 25.8) (1988-94) -13.7 (−21.1, −5.6) (1993-94)	-0.1 (-3.1, 2.9) (1995-98)			Estimates from joinpoint regression (95% CI)	The study did not report estimates for other racial/ethnic groups.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	Black *Estimated annual percentage change in age-adjusted (1970 US standard) incidence rates*		2.3 (1.8-2.8) (1973-87)	21.5 (16.7, 26.5) (1988-91) -14.5 (−14.5, −0.9) (1991-93)	0.4 (-3.8, 4.8) (1994-98)			Estimates from joinpoint regression (95% CI)	Reporting delay data were available only after 1988.
	SEER	Black *Estimated annual percentage change in age-adjusted (1970 US standard) reporting delay-adjusted incidence rates*		2.2 (1.7-2.7) (1973-87)	20.7 (16.0, 25.6) (1988-91) -8.7 (-14.8, -2.3) (1991-93)	-2.1 (-5.6, 1.5) (1994-98)			Estimates from joinpoint regression (95% CI)	Reporting delay data were available only after 1988.
Merrill[32] 2002 11790678	SEER	White Black		 (1989)	 (1992)				+42% +35% Percentage increase of point prevalence corrected incidence	Additional information is presented in Figures 1-4 of the paper (1975-97).
Sarma[34] 2002 11828352	SEER	White Black *Age-adjusted rates*		86 124 (1986)	179 (1992) 250 (1993)				+108% increase +102% increase	Also Figure 2 (1981-98)
	SEER	White Black			 (1992)	 (1998)			-5.7%/yr -4.0%/yr (annual change in incidence yr)	

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Shao[42] 2009 19713548	SEER	White Black *Incidence rate (SD)*		220.3 (1.5) 295.4 (6.3) (1988-89)		257.9 (1.5) 434.6 (6.7) (1996-97)		234.4 (1.3) 368.0 (5.3) (2004-05)	+14.1 (P<0.001) +72.6 (P<0.001) [absolute change between 1988-89 and 2004-05 (p-values from ANOVA with linear contrast)]	
Merrill[43] 1996 8931614	SEER	White Black *Age-adjusted (1990 standard) incidence rate*		138.0 180.8 (1989)	221.6 298.5 (1992)				+61% (White) +65% (Black) [percent increase from 1989 to 1992] -16% (White) +2% (Black) [percent change from 1992 to 1993]	
Zhu[45] 2009 19505907	SEER	White Black			1.25 (1.18, 1.31) 2.32 (2.07, 2.61) (1990-94)		2.63 (2.53, 2.72) 6.33 (5.95, 6.75) (2000-04)		2.11 (1.98, 2.25) 2.73 (2.39, 3.11) [IRR of 2000-04 vs. 1990-94 (95% CI))	No data were reported for other racial or ethnic groups.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Polednak[46] 2002 12477140	SEER	Non-Hispanic White Black			151.4 179.0 (1992)	122.1 183.5 (1997)			Age-adjusted incidence rates per 100,000 US 1990 standard population	Limited to locoregional prostate cancer. Data were not reported for patients of other racial/ethnic groups.
Paltoo[83] 2004 15219802	SEER	American Indian/Alaska Native White *Change in age-adjusted rates (2000 standard)*			(1992)	(1999)			-5.75%/y (P=NS) -4.70%/y (P<0.05) [Annual percentage change]	Additional information is presented in the single Figure of the paper.
Kohler[84] 2011 21454908	SEER, NPCR, NAACCR	White Black Asian/Pacific Islander American Indian/ Alaska Native Hispanic Non-hispanic *Age-adjusted (2000 US standard) incidence rates*				(1998)		(2007)	-0.7 (P=NS) -1.0 (P=NS) -1.5 (P=NS) -2.3 (P<0.05) -2.0 (P<0.05) -0.3 (P=NS) [Average annual percentage change, 1998-2007]	
Demers[49] 1994 8203988	SEER-Detroit	White Black *Age-adjusted (1970 US standard) incidence rates*		102 141 (1988)	178 218 (1991)				NR	Additional information is presented in Figure 1 of the paper.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Potosky[57] 1995 7530782	SEER-Medicare	White Black *Age-adjusted (1970 US standard) incidence rate*		840 1137 (1987)	1310 1848 (1991)				NR	Men ≥65 yr old
Danley[76] 1995 8580296	LAC/USC CSP	White – non Hispanic Black/African American Asian *Percent annual change in age-adjusted incidence rates (95% CI) using 1976 as the baseline*		2.7 (2.3, 3.1) 0.7 (-0.1, 1.6) -0.7 (-2.8, 1.3) 0.2 (-0.8, 1.2) (1988)					P<0.001 (difference between racial/ethnic groups for the linear trend over 1976-88)	Additional information is presented in Figure 1 of the paper.
Tumor characteristics – Stage										
Hankey[17] 1999 10379964	SEER	Localized/regional Distant Unstaged *Annual percentage change of the age-standardized (1970 US standard) incidence rates*	NA NA -15.3 (−20.1, -9.3) (1973-80)	3.3 (2.4, 4.2) (1973-88) 1.4 (0.5, 2.3) (1977-86)	18.7 (10.6, 27.4) (1988-92) -1.3 (-4.1, 1.6) (1986-92) 17.9 (14.8, 21.0) (1980-92)	-9.8 (-15.5, -3.9) (1992-95) -17.9 (−20.8, −14.9) (1991-95) -22.5 (−32.7, −10.8) (1992-95)			Annual percentage changes calculated from joinpoint regression (95% CI).	Additional information is presented in Figure 3 of the paper.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Chu[18] 2003 12627516	SEER	50-59 60-69 70-79 80-84 ≥85 *Age-adjusted (1990 US standard) incidence rates of distant disease by age at diagnosis [White]*		9.5 47.9 140.6 238.9 229.1 (1986)		4.7 21.5 41.2 64.0 76.6 (1999)			-50.5 -55.1 -70.7 -73.2 -66.6 (Relative change %)	Data were not reported for other racial/ethnic groups.
	SEER	50-59 60-69 70-79 80-84 ≥85 *Age-adjusted (1990 US standard) incidence rates of distant disease by age at diagnosis [Black]*		28.5 117.6 274.1 260.4 264.3 (1986)		18.8 50.1 89.4 147.2 129.3 (1999)			-34.0 -57.4 -67.4 -43.5 -51.5 (Relative change %)	Data were not reported for other racial/ethnic groups.
Dennis[23] 2000 10679753	SEER	Localized Distant Unknown *Age-adjusted incidence rates*		191.1 50.3 28.0 (1985)	461.5 42.0 91.5 (1992)				NR	
	SEER	Localized Distant Unknown *Annual relative percentage change in age-adjusted incidence rate*	+3.2% +2.5% -1.4% (1973-85)	+14.3% -1.7% +20.5% (1985-92)	-6.2% -15.9% -24.2% (1992-1996)				P<0.001 (logistic regression with a linear term for time)	

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	Positive lymph nodes Negative lymph nodes Not examined *Age-adjusted incidence rates*		12.7 53.0 265.3 (1988)	15.5 146.4 433.1 (1992)				NR	
	SEER	Positive lymph nodes Negative lymph nodes Not examined *Annual relative percentage change in age-adjusted incidence rate (percentage change during the corresponding period)*		+9.7% (+45%) +33.6% (+219%) +14.2% (+70%) (1988-92)	-23.2% (−65%) -8.2% (−29%) -9.2% (NR) (1992-1996)				P<0.001 (logistic regression with a linear term for time)	
Merrill[20] 1997 9229202	SEER	Localized Regional Distant Unstaged *Age-standardized (1970 US standard) rates [White]*	53 10 14 4 (1982)	62 15 15 11 (1987)	110 36 11 25 (1992)					White men only
	SEER	Localized Regional Distant Unstaged *Age-standardized (1970 US standard) rates [Black]*	77 13 35 6 (1982)	78 17 34 18 (1987)	138 37 32 43 (1992)				NR	Black men only

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Stephenson[26] 2002 12109343	SEER	Distant *Age-adjusted incidence rate*			16.2 (1990)	6.3 (1997)			-61% (relative change in incidence rate of distant disease)	Additional data are presented in Figures 1-5 and 14 of the paper.
Escobedo[27] 2004 15542264	SEER-Connecticut	Localized/regional Distant Unstaged *Incidence rates*	63.0 36.0 10.9 (1973-88)		158.1 26.1 23.5 (1989-98)				NR	African American men only
	SEER-Iowa	Localized/regional Distant Unstaged *Incidence rates*	66.3 36.7 7.5 (1973-88)		141.1 28.6 24.1 (1989-98)				NR	African American men only
	SEER-New Mexico	Localized/regional Distant Unstaged *Incidence rates*	69.4 22.6 8.4 (1973-88)		115.9 15.2 5.0 (1989-98)				NR	African American men only
Sarma[34] 2002 11828352	SEER	Localized disease *Rates among White individuals*		62.6 (1987)		117.7 (1998)			NR	Additional data are presented in Figure 3 of the paper (1981-98). Numerical data for localized disease were reported separately for Black and White individuals.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	Localized disease *Rates among Black individuals*		78.8 (1987)		190.2 (1998)			NR	Additional data are presented in Figure 3 of the paper (1981-98). Numerical data for localized disease were reported separately for Black and White individuals.
	SEER	Distant disease *Rates per 100,000 among White and Black individuals combined*		14.9 (1985)		6.6 (1995)			NR	Additional data are presented in Figure 3 of the paper (1981-98). Numerical data for distant disease were reported in aggregate for Black and White individuals.
Lu-Yao[35] 1994 7905093	SEER	Local Regional Distant *Annual percent change in incidence*		+5.9%/yr +10.6%/yr +0.4%/yr (1983-87)					NR	

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Shao[42] 2009 19713548	SEER	T stage 1 2 3 or 4 *Incidence rate per 100,000 (SD), age-adjusted to the 2000 US standard population in 5-yr age increments beginning at 25 yr.*		42.3 (0.6) 95.0 (1.0) 55.5 (0.7) (1988-89)		90.3 (0.9) 137.0 (1.1) 44.6 (0.6) (1996-97)		118.4 (0.9) 106.3 (0.9) 8.4 (0.2) (2004-05)	+76.1 (P<0.001) +11.2 (P<0.001) -47.1 (P<0.001) [absolute change between 1988-1989 and 2004-05 (p-values from ANOVA with linear contrast)]	
Severson[50] 1995 7707440	SEER-Detroit	Local Regional Distant Unknown *Rate*	591.1 54.3 234.2 166.7 (1982)	715.7 76.1 225.6 187.6 (1987)	1435.1 129.4 176.6 424.5 (1992)				NR	Age ≥75 yr Additional data (1973-92) are presented in Table 2 of the paper
Schwartz[52] 1999 10197854	SEER-Detroit	Local disease *Age-adjusted incidence rate*	56.5 (1982)		167.9 (1992)	129.6 (1996)			NR	Additional information is provided in Figure 1 of the paper

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Schwartz[85] 1996 8826949	SEER – Detroit	Local Regional Distant Unknown Change in incidence rate *[White]*	(1973)		(1994)				7.4 (7.1, 7.6) 11.3 (10.5, 12.1) -0.5 (-1, −0.03) 8.8 (8.2, 9.4) [Annual percentage change and corresponding 95% CI]	Only white participants.
	SEER – Detroit	Local Regional Distant Unknown Change in incidence rate *[White]*		(1989)	(1994)				9.1 (7.8, 10.3) 4.6 (1.5, 7.8) -15.4 (-18.4, 12.2) 13.4 (10.5, 16.3) [Annual percentage change and corresponding 95% CI]	Only white participants.
	SEER – Detroit	Local Regional Distant Unknown Change in incidence rate *[Black]*	(1973)		(1994)				5.6 (5.2, 6) 9.8 (8.5, 11.1) -0.6 (-1.3, 0) 11 (10, 12.2) [Annual percentage change and corresponding 95% CI]	Only black participants.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER – Detroit	Local Regional Distant Unknown Change in incidence rate *[Black]*		(1989)	(1994)				12.9 (10.5, 15.4) 12.5 (6.8, 18.6) -13.1 (-17.0, −9) 26.3 (21, 31.7) [Annual percentage change and corresponding 95% CI]	Only black participants.
Newcomer[55] 1997 9302136	SEER, Seattle-Puget Sound	Localized Regional Distant Unknown Age-adjusted rate (1970 US standard)	(1984)		(1991)				+120% +188% NR +120% Change in incidence rate, 1984 vs. 1991	Additional information is presented in Figure 2 of the paper.
	SEER, Seattle-Puget Sound	Distant Age-adjusted rate (1970 US standard)		42 (1986)	18 (1991)				P<0.001 (-60% change in incidence rate, 1986 vs. 1991)	Additional information is presented in Figure 2 of the paper.
Stephenson[56] 1996 8608513	SEER-Utah	Distant *Rate*	14.0 (1984)	15.0 (1989)	9.3 (1993)				NR	Additional data are presented in Figure 2 of the paper (1984-93)

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Danley[76] 1995 8580296	LAC/USC CSP	Localized Regional Metastatic Unstaged *Percent annual change of the stage-specific incidence rate [Non-Hispanic White]*	1.8 (1.4, 2.3) 11.3 (9.9, 12.6) 1.6 (0.7, 2.4) 3.2 (2.1, 4.4) (1976-88)						P<0.001 (for difference between stage groups for the linear trend over yr, 1976-88)	Additional information is presented in Figure 2 of the paper.
	LAC/USC CSP	Localized Regional Metastatic Unstaged *Percent annual change of the stage-specific incidence rate [Hispanic White]*	-1.0 (−2.3, 0.4) 8.3 (4.3, 12.5) 0.4 (−1.9, 2.6) 1.7 (−1.8, 5.3) (1976-88)						P<0.001 (for difference between stage groups for the linear trend over yr, 1976-88)	Additional information is presented in Figure 2 of the paper.
	LAC/USC CSP	Localized Regional Metastatic Unstaged *Percent annual change of the stage-specific incidence rate [Black]*	-0.5 (−1.6, 0.6) 7.7 (4.5, 10.9) 1.4 (−0.4, 3.1) 1.2 (−1.3, 3.8) (1976-88)						P<0.001 (for difference between stage groups for the linear trend over yr, 1976-88)	Additional information is presented in Figure 2 of the paper.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	LAC/USC CSP	Localized Regional Metastatic Unstaged *Percent annual change of the stage-specific incidence rate [Asian]*	-1.7 (−4.2, 1) 8.9 (1.3, 11.7) -0.9 (−5.5, 4) -4.6 (−10.7, 2.0) (1976-88)						P<0.05 (for difference between stage groups for the linear trend over yr, 1976-88)	Additional information is presented in Figure 2 of the paper.
Tumor characteristics – Gleason score and histological grade										
Hankey[17] 1999 10379964	SEER	Well diff. Moderately diff. Poorly diff./undiff. Unknown *Annual percentage change of the age-standardized (1970 US standard) incidence rates*	 9.1 (7.6, 10.6) (1973-88) 5.8 (5.3, 6.3) (1973-90) -6.3 (-8.1, -4.5) (1973-88)	NA 3.2 (2.4, 4.0) (1973-92) 26.9 (19.7, 34.5) (1988-92) 18.3 (8.0, 29.6) (1990-92) 9.5 (-5.4, 26.7) (1988-92)	 -20.4 (−27.3, −12.9) (1992-95) -9.3 (-13.7, -4.7) (1992-95) -14.7 (−17.8, −11.4) (1992-95) -13.4 (−25.2, 0.3) (1992-95)				Annual percentage changes calculated from joinpoint regression (95% CI).	Additional information is presented in Figure 4 of the paper.

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Farkas[21] 1998 9730458	SEER	White Black *[Rate of increase in moderately differentiated disease incidence per 100,000 persons]*			Peak, +0.56 (1991) Start of decline, −0.20 (1993) Peak, +0.47 (1991) Start of decline, −0.15 (1994)				NR	
Shao[42] 2009 19713548	SEER	2-4 5-7 8-10 *Per 100,000 (SD), age-adjusted to the 2000 US standard population in 5-yr age increments beginning at 25 yr.*		56.2 (0.7) 88.8 (0.9) 47.5 (0.7) (1988-89)		24.9 (0.5) 167.3 (1.2) 50.7 (0.7) (1996-97)		2.3 (0.1) 193.3 (1.2) 38.3 (0.5) (2004-05)	-53.9 (P<0.001) 104.5 (P<0.001) -9.2 (P<0.001) [absolute change between 1988-1989 and 2004-05 (p-values from ANOVA with linear contrast)]	
Schwartz[52] 1999 10197854	SEER-Detroit	Well diff. Moderately diff. Poorly diff. *Age-adjusted incidence rate*	28.2 22.8 NR (1982)		NR 122.9 NR (1993)	12.2 ~100 NR (1996)			NR	Additional information is provided in Figure 2 of the paper

Appendix Table C1.2. Trends in prostate cancer incidence rates (per 100,000 unless otherwise stated) (continued)

Author Year PMID	Database	Groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Stephenson[56] 1996 8608513	SEER-Utah	2-4 5-7 8-10 *Rate per 100,000*	32.7 30.2 21.4 (1984)	30.9 43.9 20.3 (1989)	36.0 96.5 30.4 (1993)				P = 0.0105 [rank-sum test comparing age-adjusted rates between 1990-94 and 1984-89 for all 3 groups of Gleason score] P<0.001 [ANCOVA comparing the age-adjusted rate and slope of Gleason 5-7 tumors vs. the other groups, i.e. 2-4 and 8-10]	Additional information is provided in Figure 3 of the paper (1984-93)

Ordering of the studies follows Appendix Table C1.1.
ANOVA, analysis of variance; NA = not applicable; NR = not reported; RR = relative rate; PMID = PubMed identification number; SD = standard deviation; yr = years.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Patient characteristics – Age										
Hankey[17] 1999 10379964	SEER	50-59 60-69 70-79 80-84 ⩾85 *Annual percentage change of the age-standardized (1970 US standard) mortality rates*	NA 0.4 (0.1, 0.6) (1969-84) 0.3 (0.1, 0.4) (1969-85) 0.4 (0.1, 0.6) (1969-85) 1.7 (1.4, 1.9) (1969-88)	0.7 (0.4, 1.0) (1969-89) 1.9 (1.1, 2.7) (1984-91) 1.9 (1.2, 2.7) (1985-92) 2.6 (1.4, 3.9) (1985-91) 4.1 (2.1, 6.1) (1988-93)	-2.4 (-4.8, −0.1) (1989-95) -4.2 (-5.7, −2.7) (1991-95) -2.7 (-3.7, −1.7) (1992-95) -1.4 (-2.9, 0.0) (1991-95) -1.1 (-6.2, 4.3) (1993-95)				Annual percentage changes calculated from joinpoint regression (95% CI).	Additional information is presented in Figure 2 of the paper.
Chu[18] 2003 12627516	SEER	50-59 60-69 70-79 80-84 ⩾85 *Age-adjusted (1990 US standard) mortality rates by age at death [White]*		7.3 51.3 188.9 404.9 612.4 (1986)		5.4 37.0 146.4 356.7 661.3 (1999)			-26% -27.9% -22.5% -11.9% +8.0% (Relative change %) P<0.001 (slope decreased of calendar period effects in 1991 based on APC analysis)	Additional information is presented in Figures 1-4 of the paper. Data were not reported for other racial/ethnic groups.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	50-59 60-69 70-79 80-84 ⩾85 *Age-adjusted (1990 US standard) mortality rates by age at death [Black]*		19.4 132.9 438.7 791.7 986.4 (1986)		17.7 115.1 393.6 818.8 1254.5 (1999)			-8.8 -13.4 -10.3 +3.4 +27.2 (Relative change %) P<0.001 (slope decreased of calendar period effects in 1991 based on APC analysis)	Additional information is presented in Figures 1-4 of the paper. Data were not reported for other racial/ethnic groups.
Merrill[24] 2000 10792091	SEER	50-59 60-69 70-79 ⩾80		Referent 0.63 (0.46, 0.87) 0.46 (0.34, 0.63) 0.34 (0.25, 0.47) (1988-89)	Referent 0.79 (0.57, 1.10) 0.57 (0.41, 0.78) 0.53 (0.39, 0.73) (1992-93)	Referent 0.76 (0.56, 1.04) 0.58 (0.43, 0.78) 0.49 (0.36, 0.66) (1994-95)			OR (95% CI) for death from prostate cancer vs. non-prostate cancer, stratified by age.	Data were reported only for White and Black patients.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
McDavid[28] 2004 15192905	SEER, NCHS	50-59 60-69 70-79 ⩾80 *Annual percentage change of the mortality rate for each time period.*	NA 0.5 (1969-84) 0.2(1969-84) 1.2 (1969-87)	0.8 (1969-90) 1.8 (1984-92) 1.7 (1984-92) NA	NA NA NA 2.9 (1987-93)	-3.6 (1990-99) -5.6 (1992-99) -4.4 (1992-99) -3.4 (1993-99)			All results were significant with P<0.05.	The reporting of specific time intervals was determined by joint-point analysis. Additional information is presented in Figure 5 of the paper.
Jani[30] 2008 18845997	SEER	40-49 ⩾80 *5-yr prostate-cancer specific mortality*	36% 33% (1979-83)	18% 17% (1984-88)	1% 16% (1989-93)	2% 17% (1994-98)			NR	Estimated from graph. Reported as survival. Also data for other age deciles.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Collin[33] 2008 18424233	SEER	55-64 65-74 ⩾75 *Annual percentage change in age-adjusted (European Standard Population) mortality rates*	0.47 (0.01, 0.93)* (1975-87) 0.41 (0.02, 0.80)* (1975-84) 1.06 (0.87, 1.24)* (1975-87)	4.27 (-3.56, 12.74) (1987-90) 1.71 (0.99, 2.45)* (1984-91) 2.80 (2.08, 3.52)* (1987-93)	-4.14 (-4.50, -3.79)* (1990-2004) -2.85 (-6.89, 1.36) (1991-94) -3.56 (-3.89, -3.24)* (1993-2002)	NA NA NA	NA -5.07 (-5.38, -4.76)* (1994-2004) -5.32 (-8.23, -2.32)* (2002-04)		Estimates (95% CI) from joinpoint regression; time intervals differ between age groups because they were determined from the regression * denotes P<0.05	
Devesa[39] 1995 7707404	SEER	35-54 55-74 ⩾75 *Age-adjusted standardized (1970 US standard) mortality rate*		1.2 56.1 392.6 (1987-91)					+0.1 (+9.1%) +6.2 (+12.4%) +51.5 (+15.1%) *Absolute [relative] change in rate, compared to 1975-79*	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Merrill[40] 2000 10647666	SEER	50-59 60-69 70-79 ⩾80 *Percentage of prostate cancer deaths out of all deaths observed among prostate cancer patients*		61% 50% 39% 30% (1988)	52% 51% 37% 30% (1992)	52% 41% 31% 26% (1995)			NR	Excluded men younger than 50 yr for this analysis. Additional information is presented in Figure 2 of the paper.
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	<60 60-79 80+ *5-year relative survival rates*	0.27 (0.21, 0.33) 0.33 (0.31, 0.36) 0.33 (0.27, 0.39) (1980-1989)		0.38 (0.28, 0.48) 0.36 (0.32, 0.39) 0.36 (0.30, 0.43) (1990-1997)				NR	Only patients with distant prostate cancer
Patient characteristics – Race/ethnicity										
Mebane[16] 1990 2258952	SEER	Black White *Age-adjusted (1970 US standard) mortality rate*	23.5 42.4 (1980)	21.9 45.8 (1985)					NR	Data presented only for Black and White individuals.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Hankey[17] 1999 10379964	SEER	White Black Other *Annual percentage change of the age-standardized (1970 US standard) mortality rates*		0.7 (0.6, 0.8) (1969-87) 1.6 (1.4, 1.7) (1969-88) NA	3.1 (1.8, 4.4) (1987-91) 3.2 (1.3, 5.2) (1987-92) 2.6 (2.0, 3.3) (1973-92)	-1.9 (−2.6, −1.1) (1991-95) -1.7 (−3.5, 0.1) (1992-95) -1.8 (−8.0, 4.9) (1992-95)			Annual percentage changes calculated from joinpoint regression (95% CI).	Additional information is presented in Figure 1 of the paper.
Chu[18] 2003 12627516	SEER	White Black *1 yr survival rates of patients diagnosed with prostate cancer by year of diagnosis*	96% 94% (1981-85)			98% 98 (1992-97)			NR	
	SEER	White Black *3 yr survival rates of patients diagnosed with prostate cancer by year of diagnosis*	85% 78% (1981-85)			94% 92% (1992-97)			NR	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	White Black *5 yr survival rates of patients diagnosed with prostate cancer by year of diagnosis*	76% 68% (1981-85)			91% 87% (1992-97)			NR	
Brawley[19] 1997 9351560	SEER	White Black *Age-adjusted standardized (1970 US standard) mortality rate*	(1973)		(1994)				Change in mortality: +18.6% among White Americans; +41.4% among Black Americans	Additional information is provided in Figures 1 and 2 of the paper
	SEER	*Proportion surviving 5 yr or longer after diagnosis*	73.9% (1981) 61.0% (1973)	87.4% (1989)					NR	
	SEER	White men only *Age-adjusted standardized (1970 standard) mortality rate*			All regions examined: 1991, 24.7; 1994, 23.8 (1991-94) Connecticut: 1991, 25.3; 1994, 23.0 (1991-94)				All regions: decrease of 0.9 deaths per 100,000 individuals per yr All regions: decrease of 2.3 deaths per 100,000 individuals per yr	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Merrill[24] 2000 10792091	SEER	White Black		Referent 0.9 (0.8, 1.04) (1988-89)	Referent 1.11 (0.98, 1.27) (1992-93)	Referent 1.06 (0.94, 1.21) (1994-95)			OR (95% CI) for death from prostate cancer vs. non-prostate cancer, comparing White and Black men.	No data were reported for patients belonging to other racial/ethnic groups.
Stephenson[26] 2002 12109343	SEER	Black *Age-adjusted mortality rate*		(1993)	(1997)			-10.9% relative change of mortality rate	Black men only. Additional information is provided in Figures 9-12 and 14 of the paper.	
Escobedo[27] 2004 15542264	SEER-Connecticut	African American men only *Age-adjusted standardized (1970 standard) mortality rate*	78.2 (65.9-90.6) (1979-86)	81.2 (65.1-97.3) (1987-90)	93.1 (82.3-103.8) (1991-98)				NR	
	SEER-Iowa	African American men only *Age-adjusted standardized (1970 standard) mortality rate*	79.5 (55.1-103.9) (1979-86)	111.1 (71.9, 105.3) (1987-90)	93.5 (69.1-117.9) (1991-98)				NR	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER-New Mexico	African American men only *Age-adjusted standardized (1970 standard) mortality rate*	104.7 (69.4-140.0) (1979-86)	62.1 (26.9-97.4) (1987-90)	47.6 (29.6-65.5) (1991-98)				NR	
Sarma[34] 2002 11828352	SEER	White Black *Age-adjusted (1970 US standard) mortality rate*	21.0 (1981) 45.8 (1981)		24.7 (1991) 56.2 (1993) (peak values for each racial group)				Since 1993, Black individuals experienced a 3% annual decrease in age-adjusted mortality rate; White individuals experienced annual decreases in age-adjusted mortality rate of 1.4% during 1991-94 and 4.7% during 1994-98.	Additional information is presented in Figure 4 of the paper.
		White vs. Black 5-yr prostate cancer mortality		+14.1% (1983-87)	+14.6% (1988-93)				Average 5-yr relative survival of White vs. Black patients	Additional information is presented in Figure 5 of the paper.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Merrill[40] 2000 10647666	SEER	White Black Other/unknown *Percentage of prostate cancer deaths out of all deaths observed among prostate cancer patients*		36% 41% 39% (1988)	35% 43% 34% (1992)	29% 34% 31% (1995)			NR	Additional information is presented in Figure 3 of the paper.
Stewart[44] 2004 15179359	SEER, NCHS	All races White White non-Hispanic White Hispanic Black American Indian/ Alaska native Asian/Pacific Islander Hispanic *Age-adjusted (2000 US standard) death rate*			38.6 35.7 34.6 24.0 78.0 19.8 16.7 23.6 (1990)		30.6 27.9 28.1 22.5 69.2 20.1 12.8 22.2 (2000)		-2.6% (P<0.05) -2.8% (P<0.05) -2.5%(P<0.05) -1.1% -1.4%(P<0.05) -1.5% -3.4%(P<0.05) -1.0% (annual percentage change, P-values for the null hypothesis that there was no change)	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Kohler[84] 2011 21454908	SEER, NPCR, NAACCR, NCHS	White Black Asian/Pacific Islander American Indian/ Alaska Native Hispanic Non-hispanic *Age-adjusted (2000 US standard) mortality rates*				(1998)		(2007)	-3.8 (P<0.05) -4.2 (P<0.05) -3.1 (P<0.05) -1.6 (P=NS) -3.8 (P<0.05) -3.8 (P<0.05) [Average annual percentage change, 1998-2007]	
Demers[51] 2001 11745285	SEER-Detroit	White Black *Age-adjusted (1970 US standard) mortality rates*		23.1 44.0 (1988)	25.4 49.5 (1992)	21.7 44.0 (1997)			P<0.001 (decreasing trend in mortality rates, 1993-98)	Additional information is presented in Figure 1 of the paper.
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	Caucasian African American *5-year relative survival rates*	0.33 (0.30, 0.36) 0.31 (0.28, 0.35) (1980-1989)		0.36 (0.32, 0.39) 0.36 (0.31, 0.41) (1990-1997)				NR	Only patients with distant prostate cancer
Godley[60] 2003 14625261	SEER-Medicare	Non-Hispanic White Black *Median survival time following diagnosis, yr*		8.4 (8.3 to 8.7) 6.8 (6.3 to 7.2) (1986-88)	Not reached 7.8 (7.3 to 8.3) (1989-91)	Not reached Not reached (1992-96)			NR	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
		Non-Hispanic White Black *5-yr survival rate following diagnosis*		0.70 (0.69 to 0.71) 0.62 (0.59 to 0.65) (1986-88)	0.77 (0.76 to 0.78) 0.68 (0.65 to 0.70) (1989-91)	0.80 (0.79 to 0.81) 0.75 (0.73 to 0.77) (1992-96)			NR	
		Non-Hispanic White Black *10-yr survival rate following diagnosis*		0.42 (0.41 to 0.43) 0.32 (0.29 to 0.34) (1986-88)	0.53 (0.51 to 0.54) 0.39 (0.35 to 0.43) (1989-91)	Not reached Not reached (1992-96)			NR	
		Non-Hispanic White Black *Kaplan-Meir survival rate at 12 mo*		0.97 (0.96 to 0.98) 0.95 (0.92 to 0.98) (1986-88)	0.98 (0.98 to 0.99) 0.93 (0.91 to 0.96) (1989-91)	0.99 (0.98 to 0.99) 0.97 (0.96 to 0.98) (1992-96)			P<0.001 (log-rank p-value comparing White with Black individuals)	
		Non-Hispanic White Black *Kaplan-Meir survival rate at 60 mo*		0.81 (0.80 to 0.83) 0.72 (0.66 to 0.79) (1986-88)	0.86 (0.85 to 0.87) 0.79 (0.75 to 0.82) (1989-91)	0.88 (0.87 to 0.89) 0.84 (0.82 to 0.87) (1992-96)			P<0.001 (log-rank p-value comparing White with Black individuals)	
		Non-Hispanic White Black *Kaplan-Meir survival rate at 120 mo*		0.57 (0.55 to 0.59) 0.46 (0.39 to 0.53) (1986-88)	0.67 (0.64 to 0.70) 0.56 (0.49 to 0.63) (1989-91)	Not reached Not reached (1992-96)			P<0.001 (log-rank p-value comparing White with Black individuals)	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Danley[76] 1995 8580296	LAC/USC CSP	White – non Hispanic Black/African American Asian *Percent annual change in age-adjusted mortality rates (95% CI) using 1976 as the baseline*		0.5 (-0.2, 1.2) 0.3 (-1.8, 2.5) 1.6 (0.0, 3.2) 2.3 (-2.9, 7.80					P>0.05 (difference between racial/ethnic groups for the linear trend over yr, 1976-88)	Additional information is presented in Figure 1 of the paper.
Patient characteristics – Comorbidities										
Merrill[24] 2000 10792091	SEER	No Yes *Multiple primary cancers*		Referent 0.36 (0.32, 0.41 (1988-89)	Referent 0.38 (0.34, 0.43) (1992-93)	Referent 0.35 (0.32, 0.40) (1994-95)			OR (95% CI) for death from prostate cancer vs. non-prostate cancer, stratified by comorbidity status	Data were reported only for White and Black patients.
Tumor characteristics – stage										
Chu[18] 2003 12627516	SEER	Local/regional Distant Unstaged *1 yr survival rates of patients diagnosed with prostate cancer by year of diagnosis [White]*	98% 84% 96% (1981-85)			100% 81% 97% (1992-97)			NR	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	Local/regional Distant Unstaged *1 yr survival rates of patients diagnosed with prostate cancer by year of diagnosis [Black]*	98% 83% 97% (1981-85)			99% 81% 97% (1992-97)			NR	
	SEER	Local/regional Distant Unstaged *3 yr survival rates of patients diagnosed with prostate cancer by year of diagnosis [White]*	93% 50% 84% (1981-85)			98% 49% 92% (1992-97)			NR	
	SEER	Local/regional Distant Unstaged *3 yr survival rates of patients diagnosed with prostate cancer [Black]*	91% 47% 83% (1981-85)			97% 48% 90% (1992-97)			NR	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	Local/regional Distant Unstaged *5 yr survival rates of patients diagnosed with prostate cancer [White]*	87% 33% 75% (1981-85)			96% 35% 87% (1992-97)			NR	
	SEER	Local/regional Distant Unstaged *5 yr survival rates of patients diagnosed with prostate cancer [Black]*	83% 30% 65% (1981-85)			93% 34% 84% (1992-97)				
Smart[78] 1997 9351557	SEER	Localized Regional Distant *[5-year relative survival rates]*	93.0% 82.2% 29.2% (1983-1987)	99.1% 93.6% 30.7% (1988-1993)					NR	
Mariotto[81] 2006 16572414	SEER	Locoregional Distant *[absolute annual increase in 5-year relative survival]*	(1975)			(1995)			+1.24 (0.06) +0.32 (0.09) Annual increase (SE), 1975-95, based on least squares regression of the survival rate over year of diagnosis	

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Merrill[24] 2000 10792091	SEER	Local Regional Distant Unknown		Referent 2.36 (2.02, 2.76) 5.21 (4.61, 5.89) 2.16 (1.85, 2.53) (1988-89)	Referent 2.49 (2.16, 2.89) 6.95 (6.14, 7.87) 2.13 (1.86, 2.43) (1992-93)	Referent 2.36 (2.05, 2.71) 6.35 (5.59, 7.21) 1.73 (1.52, 1.97) (1994-95)			OR (95% CI) for death from prostate cancer vs. non-prostate cancer, stratified by stage.	Data were reported only for White and Black patients.
Merrill[40] 2000 10647666	SEER	Local Regional Distant Unknown Percentage of prostate cancer deaths out of all deaths observed among prostate cancer patients		22% 45% 64% 41% (1988)	22% 45% 67% 36% (1992)	19% 41% 63% 32% (1995)			NR	Excluded in situ cancer cases for this analysis. Additional information is presented in Figure 4 of the paper.
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	Distant prostate cancer *[5-year relative survival rates]*	0.32 (0.30, 0.34) (1980-1989)		0.36 (0.33, 0.39) (1990-1997)				NR	Only patients with distant prostate cancer

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Tumor characteristics – grade										
Merrill[24] 2000 10792091	SEER	Well diff. Moderately diff. Poorly diff./undiff. Unknown		Referent 2.39 (2.04, 2.82) 4.09 (3.48, 4.82) 2.80 (2.35, 3.36) (1988-89)	Referent 2.03 (1.72, 2.39) 4.29 (3.63, 5.06) 2.81 (2.34, 3.38) (1992-93)	Referent 2.14 (1.80, 2.55) 4.84 (4.06, 5.78) 3.55 (2.92, 4.31) (1994-95)			OR (95% CI) for death from prostate cancer vs. non-prostate cancer, stratified by grade.	Data were reported only for White and Black patients.
Merrill[40] 2000 10647666	SEER	Well-diff. Moderately diff. Poorly diff./ undifferentiated *Percentage of prostate cancer deaths out of all deaths observed among prostate cancer patients*		16% 36% 55% 39% (1988)	16% 33% 55% 38% (1992)	12% 26% 48% 36% (1995)			NR	Additional information is presented in Figure 5 of the paper.

Appendix Table C1.3. Trends in mortality rates (per 100,000 person-yr unless otherwise stated) or survival (continued)

Author Year PMID	Database	Groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	Tumor grade I II III IV unknown *[5-year relative survival rates]*	0.47 (0.39, 0.55) 0.42 (0.37, 0.46) 0.23 (0.20, 0.26) 0.17 (0.08, 0.26) 0.32 (0.27, 0.37) (1980-1989)		0.59 (0.45, 0.73) 0.50 (0.45, 0.56) 0.27 (0.23, 0.31) 0.11 (0.00, 0.24) 0.33 (0.25, 0,40) (1990-1997)				NR	Only patients with distant prostate cancer

Ordering of the studies follows Appendix Table C1.1.
APC = age-period-cohort; CI = confidence interval; Diff = differentiated; PMID = PubMed identification number; yr = year.

Appendix Table C1.4. Patient characteristics—age

Author Year PMID	Database	Age groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Brawley[19] 1997 9351560	SEER	Whites Blacks *Median age at diagnosis*	72 70 (1980)		72 70 (1993)				NR	
Farkas[21] 1998 9730458	SEER	Whites African-Americans *Mean age (95% CI)*		72.0 (71.8 to 72.3) 70.1 (69.4 to 70.8) (1985)	69.2 (69.0 to 69.3) 67.3 (66.9 to 67.8) (1994)				NR	Figure 2 of the article presents additional information (1973-94).
Hayat[29] 2007 17227898	SEER	*Median age at diagnosis*	72 (1979-83)		71 (1989-93)		68 (1999-2003)		NR	
Jani[30] 2008 18845997	SEER	40-49 50-59 60-69 70-79 ≥80	6% 8% 31% 40% 21% (1979-83)	5% 7% 31% 42% 20% (1984-88)	7% 8% 33% 43% 15% (1989-93)	2% 14% 36% 38% 11% (1994-98)	2% 18% 34% 34% 11% (1999-2003)		P=0.68 (chi-square comparing across all periods)	
Jani[37] 2007 17505529	SEER	40-49 50-59 60-69 70-79 ≥80		5% 7% 31% 42% 20% (1984-88)	7% 8% 33% 43% 15% (1989-1993)	2% 14% 36% 38% 11% (1994-1998)	2% 18% 34% 34% 11% (1999-2003)		P=0.180 (chi-square test) RR=1.15 (CI 0.89, 1.31)	

Appendix Table C1.4. Patient characteristics—age (continued)

Author Year PMID	Database	Age groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Shao[42] 2009 19713548	SEER	White Black Total *Mean age (SD)*		72.3 (8.6) 70.8 (8.9) 72.2 (8.7) (1988-89)		69.5 (12.3) 67.2 (9.7) 69.2 (12.0) (1996-97)		67.6 (10.0) 64.4 (10. 0) 67.2 (10.1) (2004-05)	-4.7 (P<0.001) -6.4 (P<0.001) -5.0 (P<0.001) [absolute change between 1988-1989 and 2004-05 (p-values from ANOVA with linear contrast)]	Excluding age <25 yr
Polednak[46] 2002 12477140	SEER	<65 ≥65 *Age at diagnosis*			4717 (non-Hispanic White), 828 (Black) 18,630 (non-Hispanic White), 2141 (Black) (1992)	5229 (non-Hispanic White), 1101 (Black) 11,882 (non-Hispanic White), 1720 (Black) (1997)			NR	Data were not reported for patients of other racial/ethnic groups.
Severson[50] 1995 7707440	SEER-Detroit	75-79 80-84 85-89 ≥90 *Count (percentage) of patients in each age group*	1079 (47%) 739 (32%) 378 (16%) 120 (5%) (1980-84)	1265 (50%) 772 (30%) 382 (15%) 124 (5%) (1985-88)	2290 (54%) 1260 (30%) 533 (13%) 158 (4%) (1989-1992)				NR	Limited to men ≥75 yr

Appendix Table C1.4. Patient characteristics—age (continued)

Author Year PMID	Database	Age groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Schwartz[52] 1999 10197854	SEER-Detroit	0-39 40-59 60-79 80+	1% 9% 70% 21% (1982-86)	1% 8% 74% 18% (1987-91)	1% 12% 75% 13% (1992-96)				NR	
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	<60 60-79 80+	9.1% 72.5% 18.5% (1980-1989)		12.6% 75.7% 11.7% (1990-1997)				NR	
Newcomer[55] 1997 9302136	SEER, Seattle-Puget Sound	Mean age at diagnosis		71.7 (1989)	70.5 (1993)				P<0.001 (statistical test NR)	
		<65 65-75 >75	21.4% 41.7% 36.9% (1983-84)	19.5% 45.5% 34.9% (1987-88)	23.6% 46.8% 29.6% (1991-92)				NR	
Stephenson[56] 1996 8608513	SEER-Utah	<70		38% (1984-90)	42% (1991-93)				P=0.002 [Fishers' exact test] OR=1.178 (CI, NR)	
Mullins[63] 2010 20163844	SEER-Medicare	65-69 70-74 75-79 80+				29.4% 29.0% 23.1% 18.5% (1998)	30.1% 28.6% 22.4% 18.9% (2002)		P<0.05 (for the association of year with age, 1998-2002)	
Cooperberg[65] 2002 12131295	CaPSURE	<60 60-69 70-79 ≥80			16.5% 41.8% 36.1% 5.6% (1989-97)	21.7% 38.2% 32.3% 7.8% (1997-2001)			NR	

Appendix Table C1.4. Patient characteristics—age (continued)

Author Year PMID	Database	Age groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Greene[71] 2005 16194711	CaPSURE	<60 60-70 >70				23% 46% 31% (1997- 99)	28% 42% 31% (2000-03)		P=0.23 (chi-square p-value comparing 2 periods)	
Mettlin[72] 1994 8062197	NCDB	<50 50-54 55-59 60-64 65-69 70-74 75-79 80-84 ≥85		0.5% 1.5% 5.3% 12.4% 20.2% 22.7% 18.7% 12.0% 6.7% (1985)	0.5% 1.6% 4.7% 11.6% 20.8% 24.3% 19.8% 10.9% 5.8% (1990)				NR	
	NCDB	Mean age		71.6 (1985)	71.6 (1990)				NR	
Mettlin[73] 1996 8640686	NCDB	<50 50-59 60-69 70-79 80+		0.6% 6.7% 31.1% 42.8% 18.7% (1987)	0.8% 7.9% 34.9% 42.8% 13.6% (1992)				NR	
Mettlin[74] 1995 8625214	NCDB	0-29 30-39 40-49 50-59 60-69 70-79 80+		0.1% 0% 0.5% 6.8% 31.2% 42.5% 18.9% (1986-87)	0.1% 0% 0.7% 8.0% 34.9% 42.9% 13.5% (1992)				NR	
Mettlin[75] 1998 9781963	NCDB	*Median age* *Mean age* *% <65*			71 70.7 21.3% (1992)	69 68.8 29.8% (1995)			NR	

Appendix Table C1.4. Patient characteristics—age (continued)

Author Year PMID	Database	Age groups (yr)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Mettlin[91] 1999 10547563	NCDB	<60 60-64 65-69 70-74 ≥75			9% 13% 22% 25% 32% (1992)	17% 16% 23% 22% 23% (1996)			NR	
Hamilton[77] 2011 20735387	POCS, NCI	<60 60-64 65-74 ≥75				20.6% 18.0% 41.6% 19.9% (1998)	32.5% 21.5% 28.5% 17.5% (2002)		P<0.01 (difference in distribution by yr)	

Ordering of the studies follows Appendix Table C1.1.
ANOVA = analysis of variance; CI = confidence interval; NA = not applicable; NR = not reported; PMID = PubMed identification number; RR = relative risk; SD = standard deviation; yr = year.

Appendix Table C1.5. Patient characteristics—comorbidity

Author Year PMID	Database	Assessment method	Comorbidity groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Greene[71] 2005 16194711	CaPSURE	Number of comorbidities	0 1-2 ≥3				17% 52% 31% (1997-99)	15% 56% 30% (2000-03)		P=0.46 (chi-square comparing 2 periods)	
Hamilton[77] 2011 20735387	POCS, NCI	Number of comorbidities	0 ≥1				78.3% 21.7% (1998)	87.4% 12.7% (2002)		P<0.01 (difference in distribution by yr)	

Ordering of the studies follows Appendix Table C1.1.
PMID = PubMed identification number; yr = year.

Appendix Table C1.6. Patient characteristics—race/ethnicity

Author Year PMID	Database	Race/ethnicity	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Hsing[a80] 2000 10585584	SEER	White Black	24,192 (90.1%) 2664 (9.9%) (1973-77)		66,227 (90.3%) 7129 (9.7%) (1988-92)				NR	No data were reported for other ethnicities.
Clegg[31] 2002 12230422	SEER	The study provided a cross-table of the proportion of patients with localized/regional, distant or unknown cancer stage, by race and period of study (1975-87 and 1988-97). The racial groups considered were non-Hispanic White, Hispanic White, African American, American Indian and Alaskan native, Asian American, and Hawaiian native.	(1975-87)		(1987-97)				P<0.001 (chi-square test for stage x race x time period)	Hispanic Whites, non-Hispanic Whites and Asian Americans had the highest proportion of cases diagnosed at localized/regional stage and the lowest proportion of cases diagnosed at distant stage. Conversely, African Americans had the lowest proportion of cases diagnosed at localized/regional stage and the highest proportion diagnosed at distant stage. These differences were generally consistent across time periods.

[a] Incidence rates and patient characteristics in this study were obtained from SEER data and were extracted in our summary tables; mortality rates were obtained from published sources and were not extracted.

Appendix Table C1.6. Patient characteristics—race/ethnicity (continued)

Author Year PMID	Database	Race/ethnicity	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Jani[37] 2007 17505529	SEER	Caucasian African-American Other		86.5% 10.0% 3.5% (1984-88)	85.0% 9.6% 5.3% (1989-1993)	80.0% 12.2% 7.5% (1994-1998)	80.3% 11.7% 7.9% (1999-2003)		P=0.785 (chi-square test) RR=1.45 (CI 0.73, 1.81)	Data were only reported for White and Black individuals.
Polednak[46] 2002 12477140	SEER	Non-Hispanic White Black Count (percentage) of men in each racial/ethnic group			23,347 (88.7%) 2969 (11.3%) (1992)	17,111 (85.8%) 2821 (14.2%) (1997)			NR	Data were not reported for patients of other racial/ethnic groups.
Demers[49] 1994 8203988	SEER-Detroit	White Black		74% 26% (1987)	77% 23% (1991)				P<0.001 (Mantel-Haenszel test for the linear association between year of diagnosis and race)	
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	Caucasian African American	72.8% 27.2% (1980-1989)		73.7% 26.3% (1990-1997)				NR	
Schwartz[52] 1999 10197854	SEER-Detroit	White Black Other Unknown	73% 27% 1% 1% (1982-86)	74% 24% 1% 2% (1987-91)	69% 25% 1% 6% (1992-96)				NR	

Appendix Table C1.6. Patient characteristics—race/ethnicity (continued)

Author Year PMID	Database	Race/ethnicity	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Gilliland[53] 1996 8722215	SEER-New Mexico	White, non-Hispanic Hispanic Black American Indian Other/unknown	71.7% 22.1% <1% 3.8% 1.5% (1983-84)	69.4% 24.4% 1.9% 4.1% <1% (1987-88)	76.8% 18.9% 1.5% 2.7% <1% (1991-92)				NR	
Klabunde[58] 1998 9749657	SEER, Medicare	White Black		91.7% 8.3% (1987)	91.4% 8.6% (1992)				NR	
Zeliadt[61] 2004 15596192	SEER, Medicare	White African-American			91.0% 9.0% (1991-1993)	89.3% 10.7% (1997-1999)			NR	
Carpenter[62] 2010 20333462	SEER, Medicare	White Black			85.8% 14.2% (1994)	82.9% 17.1% (1997)	83.5% 16.5% (2002)		P=0.043 (chi-square test comparing yr of diagnosis between White and Black men, for all yr from 1994 to 2002)	
Mullins[63] 2010 20163844	SEER-Medicare	Non-Hispanic White African American Hispanic White				80.6% 12.2% 7.2% (1998)	82.2% 10.7% 7.1% (2002)		NR	
Cooperberg[65] 2002 12131295	CaPSURE	White – non Hispanic Black/African American Hispanic Other			84.9% 11.1% 2.1% 1.9% (1989-97)	86.8% 8.8% 1.6% 2.9% (1997-2001)			NR	

Appendix Table C1.6. Patient characteristics—race/ethnicity (continued)

Author Year PMID	Database	Race/ethnicity	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Greene[71] 2005 16194711	CaPSURE	White – non Hispanic Black/African American Asian Hispanic Alaska native/American native Other				83% 11% 2% 3% 0% 1% (1997- 99)	91% 5% 1% 1% <1% 1% (2000-03)		P<0.001 (chi-square comparing 2 periods)	
Mettlin[74] 1995 8625214	NCDB	Non-Hispanic White Hispanic African American Native American Asian Unknown		85.7% 1.6% 7.6% 0.1% 0.8% 4.2% (1986-87)	85.1% 2.1% 8.1% 0.1% 1.1% 3.5% (1992)				NR	
Mettlin[75] 1998 9781963	NCDB	White – non Hispanic Black/African American Hispanic Other			87.7% 8.8% 0.8% 2.7% (1992)	83.3% 11.8% 1.6% 3.3% (1995)			NR	
Hamilton[77] 2011 20735387	POCS, NCI	White African American Hispanic				78.4% 14.0% 7.6% (1998)	83.2% 12.5% 4.3% (2002)		P<0.01 (difference in distribution by yr)	
White[92] 2011 21351084	Texas Cancer Registry	Non-Hispanic white Non-Hispanic black Hispanic Asian/Pacific Islander				76.2% 11.8% 11.5% 0.5% (1997)	73.5% 12.2% 13.5% 0.8% (2002)		Reference P=0.148 P<0.001 P<0.001 [chi-square test for the racial distribution over time]	

Ordering of the studies follows Appendix Table C1.1.
NA = not applicable; NR = not reported; PMID = PubMed identification number; RR = relative risk; yr = year.

Appendix Table C1.7. Tumor characteristics—stage

Author Year PMID	Database	Stage groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Mebane[16] 1990 2258952	SEER	Localized disease Regional disease Distant disease [White patients]	64.4% 11.3% 18.7% (1980)	59.2% 11.9% 18.9% (1985)					NR	Data presented only for Black and White individuals.
	SEER	Localized disease Regional disease Distant disease [Black patients]	59.2% 9.4% 29.2% (1980)	52.2% 12.0% 26% (1985)					NR	Data presented only for Black and White individuals.
Dennis[23] 2000 10679753	SEER	Positive nodes Not examined Lymph node status at diagnosis		4% 80% (1988)	2.6% (1992) 75% (1994)	1.3% NR (1996)			NR	
Clegg[31] 2002 12230422	SEER	The study provided a cross-table of the proportion of patients with localized/regional, distant or unknown cancer stage, by race and period of study (1975-87 and 1988-97)							P<0.001 (chi-square test for stage x race x time period)	The cross-table suggests that the proportion of patients with distant disease decreased and the proportion of patients with localized/regional disease increased over time, across all racial/ethnic groups studied.
Jani[37] 2007 17505529	SEER	AJCC stage 0/I II III		53% 20% 26% (1984-88)	41% 25% 33% (1989-93)	53% 21% 25% (1994-98)	65% 22% 13% (1999-2003)		P=0.025 (chi-square test) RR=0.86 (CI 0.70, 0.92)	Excluded AJCC stage IV.

Appendix Table C1.7. Tumor characteristics—stage (continued)

Author Year PMID	Database	Stage groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Polednak[46] 2002 12477140	SEER	Local/regional Distant			18,470 (Non-Hispanic White), 2094 (Black) 1490 (Non-Hispanic White), 380 (Black) (1992)	15,034 (Non-Hispanic White), 2406 (Black) 749 (Non-Hispanic White), 221 (Black) (1997)			NR	Data were not reported for patients of other racial/ethnic groups.
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	Caucasian African American *% men with distant disease by race/ethnicity*	15.6% 24.4% (1980-89)		4.9% 8.9% (1990-97)				NR	Data not reported for patients belonging to other racial/ethnic groups
	SEER-Detroit	<60 60-79 80+ *% men with distant disease by age group*	18.5% 17.7% 18.7% (1980-89)		4.0% 5.5% 11.2% (1990-97)				NR	
	SEER-Detroit	Tumor grade I II III IV Unknown *% men with distant disease by tumor grade*	 5.1% 16.3% 35.9% 36.6% 23.2% (1980-89)		 1.8% 3.2% 15.0% 18.5% 9.6% (1990-97)				NR	

Appendix Table C1.7. Tumor characteristics—stage (continued)

Author Year PMID	Database	Stage groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Schwartz[52] 1999 10197854	SEER-Detroit	Local Regional Distant Unknown	60% 7% 20% 13% (1982-86)	65% 9% 13% 13% (1987-91)	72% 9% 5% 13% (1992-96)				NR	
Stephenson[56] 1996 8608513	SEER-Utah	Distant Percent of new cases	13.9% (1984)	13.3% (1989)	4.8% (1993)				NR	Data were only presented for the subgroup of men with distant disease.
Carpenter[62] 2010 20333462	SEER-Medicare	OR (95% CI) for advanced versus early stage prostate cancer, 1994 is the baseline			1 (1994)	0.65 (0.55 to 0.76) (1997)	0.46 (0.38 to 0.56) (2002)		Multivariable logistic regression comparing stage at diagnosis by yr; the model included the following covariates: age at diagnosis, marital status, diagnosis yr, comorbidity score, median household income, receipt of surgery/related procedures, and comorbidity. Overall P<0.001, for all yr between 1994 and 2002.	

Appendix Table C1.7. Tumor characteristics—stage (continued)

Author Year PMID	Database	Stage groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Mullins[63] 2010 20163844	SEER-Medicare	OR for diagnosis with distant vs. in situ, local or regional disease, using 1998 as the reference year.				Reference year (1998)	OR=0.76 (0.69, 0.83) (2002)		P<0.01 OR adjusted for marital status, urban/rural living area, state-buy-in, PSA test prior to diagnosis, census tract median household income, SEER registry.	Analyses were limited to patients with available stage information. Additional information on interactions of patient age with time trends are presented in Table 6 of the paper.
Cooperberg[65] 2002 12131295	CaPSURE	T stage T1 T2 T3 T4			21.6% 68.6% 9.1% 0.7% (1989-97)	45.4% 50.2% 3.7% 0.5% (1997-2001)			NR	
Cooperberg[4] 2003 14610406	CaPSURE	T stage T1 T2a T2b T3-4			16.9% 48.2% 23.0% 11.8% (1989-92)	30.6% 36.0% 25.8% 7.5% (1996-99)	49.4% 27.2% 20.0% 3.5% (1999-2002)		P<0.001 (Mantel-Haenszel test for trend)	
Cooperberg[67] 2004 15169800	CaPSURE	T2a T1c T1b T1a			74.6% 15.2% 5.4% 4.9% (1989-92)	60.2% 35.9% 2.3% 1.6% (1996-98)	36.2% 61.7% 0.5% 1.6% (1999-2001)		NR	Low-risk prostate cancer patients only (n=1990)

Appendix Table C1.7. Tumor characteristics—stage (continued)

Author Year PMID	Database	Stage groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Cooperberg[68] 2007 17644125	CaPSURE	T stage 1a 1b 1c 2a			 3.5% 4.1% 29.9% 62.5% (1990-94)	 1.6% 1.1% 49.8% 47.4% (1995-99)	 <1% <1% 73.3% 25.3% (2002-03)	 <1% <1% 78.3% 20.7% (2004-06)	P<0.001 (“trend in distribution of each risk characteristic”)	Trend data reported only for “low-risk” prostate cancer (PSA≤10 ng/ml, Gleason score ≤6 and clinical stage T1 or T2a).
Cooperberg[90] 2008 18369637	CaPSURE	T stage T1 T2a/b T2c T3a			 12.7% 20.7% 56.7% 10.0% (1990-94)	 9.3% 14.6% 64.8% 11.4% (1995-99)	 23.4% 10.9% 61.2% 4.6% (2002-03)	 26.1% 10.4% 60.7% 2.8% (2004-07)	P<0.001 (Cuzick test for trend)	High risk localized prostate cancer only
Greene[71] 2005 16194711	CaPSURE	T stage 1 2 3				 43% 54% 2% (1997- 99)	 58% 41% 1% (2000-03)		P<0.001 (chi-square comparing 2 periods)	
Mettlin[72] 1994 8062197	NCDB	0-II III IV		67.3% 13.0% 19.7% (1985)	64.7% 16.2% 19.1% (1990)				NR	
Mettlin[73] 1996 8640686	NCDB	Stage 0-I Stage II Stage II Stage IV		44.2% 20.2% 14.6% 20.9% (1987)	29.4% 39.9% 18.4% 12.4% (1992)					
Mettlin[74] 1995 8625214	NCDB	pAJCC/cAJCC stage 0 I II III IV Unknown		 1.9% 18.8% 9.5% 6.7% 10.6% 52.6% (1986-87)	 4.7% 18.9% 33.1% 15.1% 10.4% 17.8% (1992)				NR	

Appendix Table C1.7. Tumor characteristics—stage (continued)

Author Year PMID	Database	Stage groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Mettlin[75] 1998 9781963	NCDB	Stage 0 I II III IV Localized disease (0-II)			 5.4% 22.8% 41.1% 18.0% 12.6% 69.3% (1992)	 2.0% 24.9% 49.7% 13.0% 10.3% 76.7% (1995)			NR	
Mettlin[91] 1999 10547563	NCDB	AJCC stage 0 I II III IV Unknown			 4.5% 19% 34% 15% 10% 18% (1992)	 2% 23.5% 46% 12% 8% 8% (1996)			NR	
Danley[76] 1995 8580296	LAC/USC CSP	Localized Regional/Metastatic Percentage of cases by yr of diagnosis	71% 29% (1981-84)	66% 34% (1985-88)					1.03 (0.96, 1.10) 1.29 (1.21, 1.37) (OR and 95% CIs using 1976-80 as the baseline period) P<0.001 for trend over time	

Ordering of the studies follows Appendix Table C1.1.

AJCC = American Joint Committee on Cancer; cAJC = clinical stage according to American Joint Committee on Cancer guidelines; CI = confidence interval; NA = not applicable; NR = not reported; OR = odds ratio; pAJCC = pathological stage according to American Joint Committee on Cancer guidelines; PMID = PubMed identification number; PSA = prostate-specific antigen; SD = standard deviation.

Appendix Table C1.8. Tumor characteristics—Gleason score

Author Year PMID	Database	Gleason score groups	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Perrotti[22] 1998 9720554	SEER	Well diff. Moderately diff. Poorly diff.	37.9 34.9 24.4 (1980-84)		20.4 57.6 21.4 (1990-94)				P<0.001 P<0.001 P<0.001 (comparison of proportions between time periods, separately for each grade)	Additional information is presented in Figure 1 of the paper (1974-1994).
Jani[30] 2008 18845997	SEER	Well diff. Moderately diff. Poorly diff.	41% 35% 24% (1979-83)	34% 42% 24% (1984-88)	24% 58% 18% (1989-93)	12% 69% 19% (1994-98)	4% 75% 21% (1999-2003)		P<0.001 (chi-square comparing across all periods)	
Jani[37] 2007 17505529	SEER	Well diff. Moderately diff. Poorly diff.		34% 41% 23% (1984-88)	24% 58% 18% (1989-1993)	12% 69% 19% (1994-1998)	4% 75% 21% (1999-2003)		P<0.001 (chi-square test) RR=1.37 (1.21, 1.50)	Excluded undifferentiated tumors.
Polednak[46] 2002 12477140	SEER	High grade [White]			20.6% (1992)	19.1% (1997)			P<0.001 (chi-square for linear trend in proportions for the period 1992-1998)	Data were reported only for non-Hispanic White and Black patients.
	SEER	High grade [Black]			24.7% (1992)	20.6% (1997)			P<0.005 (chi-square for linear trend in proportions for the period 1992-1998)	Data were reported only for non-Hispanic White and Black patients.

Appendix Table C1.8. Tumor characteristics—Gleason score (continued)

Author Year PMID	Database	Gleason score groups	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Barnholtz-Sloan[86] 2003 14575366	SEER-Detroit	Tumor grade I II III IV Unknown	 31.3% 31.6% 18.9% 1.7% 16.5% (1980-89)		 14.3% 56.2% 19.3% 0.6% 9.6% (1990-97)				NR	
Schwartz[52] 1999 10197854	SEER-Detroit	Well diff. Moderately diff. Poorly diff. Undifferentiated Unknown	30.2% 28.6% 18.9% 2.3% 20.0% (1982-86)	25.5% 39.4% 19.3% 1.1% 14.7% (1987-91)	12.4% 54.4% 17.9% <1% 14.8% (1992-96)				P=0.001 (proportion of moderately differentiated tumors over time)	
Stephenson[56] 1996 8608513	SEER-Utah	2-4 5-7 8-10	33% 30% 21% (1984)	27% 39% 18% (1989)	18% 50% 16% (1993)				NR	
Cooperberg[65] 2002 12131295	CaPSURE	2-4 5-6 7 8-10			17.3% 50.6% 19.9% 12.2% (1989-97)	2.9% 62.8% 24.2% 10.1% (1997-2001)			NR	
Cooperberg[4] 2003 14610406	CaPSURE	2-4 5-6 7 8-10			26.5% 46.4% 17.4% 9.7% (1989-92)	7.5% 57.7% 23.4% 11.5% (1996-99)	1.7% 63.8% 24.7% 9.8% (1999-2002)		P=0.003 (Mantel-Haenszel test for trend)	
Cooperberg[67] 2004 15169800	CaPSURE	5-6 2-4			59.5% 40.5% (1989-92)	81.6% 18.4% (1996-98)	96.1% 3.9% (1999-2001)		NR	Low-risk prostate cancer patients only (n=1990)

Appendix Table C1.8. Tumor characteristics—Gleason score (continued)

Author Year PMID	Database	Gleason score groups	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Cooperberg[68] 2007 17644125	CaPSURE	2-4 5 6			40.8% 26.4% 32.9% (1990-94)	17.4% 24.9% 57.7% (1995-99)	1.3% 4.9% 93.8% (2002-03)	0.7% 2.3% 97.0% (2004-06)	P<0.001 ("trend in distribution of each risk characteristic")	Information on trends was reported only for patients with "low-risk" prostate cancer, defined as PSA≤10 ng/ml, Gleason score≤6 and clinical stage T1 or T2a.
Cooperberg[90] 2008 18369637	CaPSURE	2 - 6 3 + 4 4 + 3 8 - 10			60.1% 13.1% 4.3% 22.5% (1990-94)	48.5% 18.7% 9.2% 23.7% (1995-99)	33.9% 19.7% 11.8% 34.6% (2002-03)	26.1% 20.9% 12.1% 40.8% (2004-07)	P<0.001 (Cuzick test for trend)	High risk localized prostate cancer only
Greene[71] 2005 16194711	CaPSURE	2-4 5-6 7 8-10				3% 66% 24% 6% (1997-99)	1% 66% 26% 7% (2000-03)		P=0.006 (chi-square comparing 2 periods)	
Mettlin[73] 1996 8640686	NCDB	Well diff. Moderately diff. Poorly diff.		31.3% 38.6% 30.1% (1986)	19.8% 57.5% 22.8 % (1993)				NR	Limited to men with known tumor grade (the proportion of unknown grade declined over time, from 18.3% in 1986 to 10.2% in 1993).
Mettlin[75] 1998 9781963	NCDB	Well diff. Moderately diff. Poorly diff. Undifferentiated			21.8% 55.6% 21.5% 1.1% (1992)	15.8% 62.2% 21.2% 0.7% (1995)			NR	

Appendix Table C1.8. Tumor characteristics—Gleason score (continued)

Author Year PMID	Database	Gleason score groups	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2010	Statistical analysis	Notes
Hamilton[77] 2011 20735387	POCS, NCI	<6 6 7 8-10 Unknown				15.6% 31.6% 15.1% 5.4% 32.4% (1998)	4.4% 52.4% 23.5% 6.9% 12.8% (2002)		P<0.01 (difference in distribution by yr)	

Ordering of the studies follows Appendix Table C1.1.
Diff = differentiated; NA = not applicable; NR = not reported; PMID = PubMed identification number; yr = year.

Appendix Table C1.9. Tumor characteristics—PSA

Author Year PMID	Database	PSA groups (ng/ml)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Kindrick[64] 1998 9817332	CaPSURE	Median PSA at diagnosis			10.1 (1992)	9.2 (1997)			NR	Additional information in provided in the single Figure of the paper.
Cooperberg[65] 2002 12131295	CaPSURE	<4 4-10 10.01-20 >20			9.6% 46.5% 22.8% 21.1% (1989-97)	13.8% 58.8% 16.9% 10.5% (1997-2001)			NR	Additional information is presented in the single Figure of the paper, depicting a line graph of median PSA (measured between diagnosis and primary treatment) for patients with "low" or "intermediate" risk cancer.
Cooperberg[4] 2003 14610406	CaPSURE	<4 4-10 10-20 >20			10.7% 37.2% 25.2% 27.0% (1989-92)	11.2% 54.8% 19.5% 14.5% (1996-99)	12.9% 63.2% 15.8% 8.1% (1999-2002)		P<0.001 (Mantel-Haenszel test for trend)	
Cooperberg[67] 2004 15169800	CaPSURE	0-4 4-10			25.5% 74.6% (1989-92)	16.5% 83.6% (1996-98)	17.2% 82.8% (1999-2001)		NR	Low-risk prostate cancer patients only (n=1990)

Appendix Table C1.9. Tumor characteristics—PSA (continued)

Author Year PMID	Database	PSA groups (ng/ml)	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Cooperberg[68] 2007 17644125	CaPSURE	<2 2-6 6-10			9.2% 42.9% 47.8% (1990-94)	4.8% 48.3% 46.9% (1995-99)	6.4% 62.0% 31.6% (2002-03)	5.8% 65.9% 28.2% (2004-06)	P<0.001 ("trend in distribution of each risk characteristic")	Information on trends was reported only for patients with "low-risk" prostate cancer, defined as PSA≤10 ng/ml, Gleason score≤6 and clinical stage T1/T2a.
Cooperberg[111] 2008 18369637	CaPSURE	0-10 10.01-20 20.01-30 >30			34.3% 23.3% 16.8% 25.6% (1990-94)	51.8% 18.3% 13.2% 16.7% (1995-99)	58.5% 14.8% 10.9% 15.8% (2002-03)	60.6% 14.9% 10.7% 13.8% (2004-07)	P<0.001 (Cuzick test for trend)	High risk localized prostate cancer only
Greene[71] 2005 16194711	CaPSURE	≤4 4.1-10 10.1-20 >20				15% 60% 18% 7% (1997-99)	16% 67% 13% 4% (2000-03)		P=0.008 (chi-square comparing 2 periods)	
Hamilton[77] 2011 20735387	POCS, NCI	0-4.0 4.1-9.9 10.0-19.9 20.0-100.0 >100.0 unknown				9.1% 53.5% 19.3% 6.2% 0.1% 11.8% (1998)	14.4% 51.1% 12.0% 6.2% 1.1% 15.2% (2002)		P<0.01 (difference in distribution by yr)	

Ordering of the studies follows Appendix Table C1.1.
NA = not applicable; NR = not reported; PMID = PubMed identification number; PSA = prostate specific antigen.

Appendix Table C1.10. Diagnostic strategies—biopsy frequency

Author Year PMID	Database	Frequency groups	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Schwartz[52] 1999 10197854	SEER-Detroit		(1982)			(1995)			P<0.001 (Cochran-Armitage test for trend in proportion of biopsies over time)	Additional information is presented in Figure 4 of the paper.
Gilliland[54] 2001 11176484	SEER-New Mexico	Biopsy only TURP Prostatectomy Other *Percentage of patients diagnosed through each procedure*	27.1% 55.6% 12.8% 4.5% (1983-84)		43.7% 11.2% 38.8% 6.3% (1992-93)				NR	Only histologically confirmed prostate cancer samples were considered.
Potosky[57] 1995 7530782	SEER-Medicare	*Age-adjusted (1970 US standard) biopsy procedure rates per 100,000 men*		685 (1986)	2600 (1991)				NR	Additional information is presented in Figure 2 of the paper.
Welch[89] 2007 17848671	SEER-Medicare	Age- and race-adjusted biopsy rates			(1993)	(2001)			P>0.2 (test for trend in the biopsy rate over time)	The authors reported that "although the biopsy rate declined over the first 3 years, there was not a statistically significant trend over the 9-year analysis period." Additional information was provided in Figure 2 of the paper.

Ordering of the studies follows Appendix Table C1.1.
NR = not reported; PMID = PubMed identification number; TURP = transurethral resection of the prostate.

Appendix Table C1.11. Diagnostic strategies— number of cores

Author Year PMID	Database	Groups by number of cores	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Shah[70] 2008 17997437	CaPSURE	Mean ±SD				7.5 ±2.1 (1997)	9.8 ±3.0 (2002)		Beta= 0.41, SE = 0.01; P<0.001 (regression of the number of removed cores on yr, from 1995 to 2004; adjusted for CaPSURE study site)	Excluded patients diagnosed with biopsies of fewer than 6 cores.

Ordering of the studies follows Appendix Table C1.1.
PMID = PubMed identification number; SD = standard deviation; SE = standard error.

Appendix Table C1.12. Gofrit 2007[99], systematic review of the effect of histopathologic grading changes

Author Year [PMID]	Gofrit 2007[99] [17997434]
Design	A structured review of studies of the "Will Rogers" phenomenon in urologic oncology
Population	All urologic cancer patients (separate results reported for prostate cancer)
Exposure	Histological reclassification (temporal change in guidelines or prevailing norms for diagnosing, staging or grading the histology of prostate cancer)
Results	Medline (15 cited studies) As reported by the authors: 1. In prostate cancer the Will Rogers phenomenon is the result of the late 1990s acceptance that Gleason scores 2-4 should not be assigned on prostate biopsy. 2. Consequently grade inflation occurred and current readings are almost 1 Gleason grade higher compared to past readings of the same biopsy. 3. The result is an illusion of improvement in grade adjusted prognosis.
Comments	As noted by the authors: 1. Comparison of contemporary results to historical controls may be biased by the Will Rogers phenomenon. 2. Ignoring the possibility of stage or grade reclassification may lead to erroneous conclusions.
AMSTAR items	
A priori design?	N
Two independent reviewers?	N
Comprehensive literature search?	N
All publication types and languages included?	N
Included and excluded studies listed?	N
Study characteristics provided?	N
Study quality assessment performed?	N
Study quality appropriately used in analysis?	N
Appropriate statistical synthesis?	NA
Publication bias assessed?	N
Conflicts of interest stated?	Y

N = no; NA = not applicable; PMID = PubMed identification number; Y = yes.

Appendix Table C1.13. System characteristics—including, differences in geographical access, insurance, physician types, etc.

Author Year PMID	Database	Characteristic reported	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Cooperberg[65] 2002 12131295	CaPSURE	***Insurance***								
		Medicare			76.5%	66.9%			NR	
		supplemental			18.8%	20.4%				
		Medicare alone			1.8%	4.8%				
		Veterans Affairs			3.0%	7.9%				
		Other			(1989-97)	(1997-2001)				
		Geographic region								
		East			48.8%	35.0%			NR	
		South			26.4%	14.2%				
		Midwest			9.0%	35.1%				
		West			15.9%	15.7%				
					(1989-97)	(1997-2001)				
		Setting								
		Community			91.2%	86.4%			NR	
		Academic			8.8%	13.6%				
					(1989-97)	(1997-2001)				
Greene[71] 2005 16194711	CaPSURE	***Insurance***								
		Medicare supplement				30%	32%		P=0.003 (chi-square comparing 2 periods)	
		Medicare				17%	12%			
		Private				52%	52%			
		Other				<1%	3%			
						(1997- 99)	(2000-03)			
Mettlin[74] 1995 8625214	NCDB	***Hospital caseload***		1.1%	0.9%				NR	
		<150 cases		19.5%	19.3%					
		150-499		38.5%	39.9%					
		500-999		30.5%	30.2%					
		1000+		10.4%	9.6%					
		Unknown size		(1986-87)	(1992)					

Appendix Table C1.13. System characteristics—including, differences in geographical access, insurance, physician types, etc. (continued)

Author Year PMID	Database	Characteristic reported	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Hamilton[77] 2011 20735387	POCS, NCI	***Insurance status***							NR	
		Private				42.8%	39.6%			
		HMO/IPA/Managed care				31.4%	34.3%			
		Medicare				12.4%	12.6%			
		Medicaid				1.2%	1.9%			
		CHAMPUS, VA, other mil.				3.2%	4.1%			
		None/unknown				9.0%	7.6%			
						(1998)	(2002)			
		Median income $ (area)							P<0.01 (difference in distribution by yr)	
		<40,000				41%	19.5%			
		40,000-75,000				47.9%	46.7%			
		>75,000				2.8%	31.9%			
		unknown				8.4%	1.9%			
						(1998)	(2002)			

Ordering of the studies follows Appendix Table C1.1.
CHAMPUS = Civilian Health and Medical Program of the Uniformed Services; HMO = health maintenance organization; IPA = independent practice association; NR = not reported; PMID = PubMed identification number; VA = Veterans Affairs; yr = year.

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Escobedo[27] 2004 15542264	SEER-Connecticut	RP TURP RT RP and RT Observation All other	6.6% 54.1% 0.0% 0.8% 1.7% 36.8% (1983-88)		21.3% 19.4% 0.6% 2.7% 1.3% 54.7% (1989-94)				NR	African American men only
	SEER-Iowa	RP TURP RT RP and RT Observation All other	11.1% 72.2% 0.0% 0.0% 1.9% 14.8% (1983-88)		25.0% 15.6% 0.0% 9.4% 1.0% 49.0% (1989-94)				NR	African American men only
	SEER-New Mexico	RP TURP RT RP+RT Observation All other	5.9% 55.9% 0.0% 2.9% 0.0% 35.3% (1983-88)		28.8% 15.0% 0.0% 7.5% 1.3% 47.5% (1989-94)				NR	African American men only

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Harlan[36] 1995 7799048	SEER	RP RT Other (including, ADT or careful observation with therapy reserved for clinical progression) *Age-adjusted proportions for initial Tx (within 4 mo of diagnosis)*	11.0% 27.0% NR (1984)		32.3% 29.7% NR (1991)				P<0.001 (chi-square test)	Additional information is presented in Figures 1-4 of the paper (1984-91).
Miller[41] 2006[a] 16912266	SEER	EM/ADT RP RT [higher-risk]					20.0% 43.1% 36.9% (2001)		P=0.042 (overall chi-square test for 2000-02)	Higher-risk patients were those with poorly diff. tumors regardless of age or those with moderately diff. tumors aged <70 yr.

[a] The study reported data for the period 2000-02, which fits entirely in one of our table's 5 yr bins. We extracted data for the midpoint of the study period (i.e., 2001) along with the p-value from tests comparing the frequency of treatments across all 3 study years. Readers are referred to the full text of the paper for additional information.

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	SEER	EM/ADT RP RT [lower-risk]		NR 17% 31% (1988-90)			43.6% 10.6% 45.9% (2001)		P=0.008 (overall chi-square test for 2000-02)	The historical cohort (1988-90) was used for comparison of treatment patterns among lower-risk patients only. Lower-risk patients were those with well-diff. tumors regardless of age or those with moderately diff. tumors aged ⩾70 yr.

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Abdollah[82] 2011 20965646	SEER	Observation RT RP		13.8% 28.8% 57.4% (1988)	19.6% 30.7% 49.8% (1992)	25.0% 32.4% 42.6% (1997)	24.1% 40.5% 35.5% (2002)	23.5% 38.1% 38.3% (2006)	"all" P<0.001 [chi-square trend tests for treatments over time]	Clinically localized prostate cancer. The authors stated that they could not account for adjuvant or salvage therapies, such as oral ADT or palliative chemotherapy, due to the unavailability of such data in the SEER database.
Underwood[48] 2004 15017208	SEER	Definitive therapy (defined as any Tx other than ADT/EM)			(1992)	(1999)			Logistic regression for trends in definitive Tx: 1. Racial/ethnic disparities improved with time in Hispanic men but less so in Black men. 2. In 1992 Hispanic men with moderately/poorly differentiated cancers were less likely to receive definitive	A cross-table of odds ratios for ethnicity and Tx year (1992-99) was provided, stratified by tumor grade (well-moderately-poorly differentiated); i.e. a 2x3x3 table. Data were not available for assessing ADT and EM

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
									therapy than White men (OR 0.62 and 0.50 for moderately and poorly differentiated tumors respectively, P<0.001 vs. White men), whereas by 1999 the odds that Hispanic men would receive definitive therapy were not significantly different from those of White men (p=0.20 and 0.96 for moderate and poorly differentiated. cancers, respectively). 3. A disparity in the use of definitive therapy by Black men with moderately or poorly differentiated cancers compared to White men persisted in 1999 (OR 0.60 and 0.45, respectively,	separately.

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
									p<0.001 for definitive therapy in each yr compared to White men). This disparity was less profound than it had been in 1992 (OR 0.56 and 0.38, respectively).	
Underwood[47] 2005 15612083	SEER	ADT/EM			(1992)	(1999)			Logistic regression for trends in Tx P<0.001	"the utilization of ADT/EM decreased significantly over time." Additional information (1992-99) was presented in a figure in the article. Data were not available for assessing ADT and EM separately.
Klabunde[58] 1998 9749657	SEER, Medicare	RP RT Conservative Tx (not radiotherapy or surgery)		(1986)	(1993)				RP: + 58.5% RT: +39.2% Conservative Tx: -38.4% (change in the proportion of men receiving each Tx over the study period, 1986-93).	White men only

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
		RP RT Conservative Tx (not radiotherapy or surgery)		(1986)	(1993)				RP: + 63.5% RT: +73.5% Conservative Tx: -37.4% (change in the proportion of men receiving each Tx over the study period, 1986-93).	Black men only
		Baseline yr = 1986 OR (95% CI) receiving aggressive Tx versus non aggressive Tx		1.12 (1.00 to 1.24) (1987)	1.39 (1.27 to 1.53) (1992)				Logistic regression predicting Tx received based on diagnosis yr; adjusted for age, socioeconomic status, SEER registry, tumor grade, comorbidity score, and the race x TURP interaction	

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Warren[87] 2008 18544740	SEER-Medicare	Cancer-related surgery RT Chemotherapy Hospitalizations for reasons other than cancer surgery *[proportions calculated among patients with claims for specific cancer services]*			54% 33% 4% 35% (1991)		24% 47% 7% 24% (2002)		-2.52 (-2.97, -2.07), P<0.001 1.40 (0.95, 1.84), P<0.001 0.15 (-0.03, 0.35), P=0.09 -0.79 (-1.03, -0.56), P<0.001 [annual % rate of decline (slopes from linear regression of the proportion of men receiving each treatment over time]	All cancer stages; some patients had claims for more than one services.
		No initial cancer-related surgery, radiation therapy, or chemotherapy ("expectant management")			21% (1991)		30% (2002)		NR	All cancer stages
Gross[88] 2008 18181101	SEER-Medicare	"Definitive therapy" (BT or EBRT or RP) No "definitive therapy" *[White]*			81.7% 18.3% (1992-94)		77.9% 22.1% (2000-02)		NR	Proportions adjusted for age, marital status, physician visits, geographic region, cancer stage and grade and comorbid conditions

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
		"Definitive therapy" (BT or EBRT or RP) No "definitive therapy" *[Black]*			74.1% 25.9% (1992-94)		69.3% 30.7% (2000-02)		NR	Proportions adjusted for age, marital status, physician visits, geographic region, cancer stage and grade and comorbid conditions
Cooperberg[4] 2003 14610406	CaPSURE	BT EBRT RP PADT WW [Low risk]			3.7% 15.3% 62.6% 4.6% 13.7% (1989-92)	5.7% 10.0% 54.6% 16.7% 12.9% (1996-98)	18.4% 6.8% 51.9% 14.2% 8.3% (1999-2001)		NR	PSA≤10.0 ng/ml, Gleason score≤6, and clinical stage T1 or T2a; excluded patients undergoing cryotherapy
		BT EBRT RP PADT WW [Intermediate risk]			3.3% 22.5% 55.3% 8.9% 10.0% (1989-92)	5.6% 18.0% 49.4% 21.3% 5.6% (1996-98)	11.8% 19.1% 45.0% 19.7% 4.5% (1999-2001)		NR	PSA≤10.1-20.0 ng/ml, Gleason score=7, or clinical stage T2a; excluded patients undergoing cryotherapy

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
		BT EBRT RP PADT WW [High risk]			3.1% 29.0% 27.3% 32.8% 7.6% (1989-92)	4.2% 19.6% 23.3% 49.7% 3.2% (1996-98)	2.4% 22.9% 22.7% 48.2% 4.8% (1999-2001)		NR	PSA⩾20.0 ng/ml, Gleason score 8-10, or clinical stage T3 or T4; excluded patients undergoing cryotherapy
Harlan[66] 2003 14532780	CaPSURE (overall)	AS/WW		7.5% (1989-91)	9.5% (1992-94)	7.9% (1995-97)	5.5% (1998-2000)		P<0.001 (Mantel-Haenszel test for trend)	
	CaPSURE (low risk)	AS/WW		7% (1989-91)	16.9% (1992-94)	11.9% (1995-97)	7.2% (1998-2000)		P=0.003 (Mantel-Haenszel test for trend) Controlling for risk group, age, comorbidity, insurance status and study site, OR for WW, compared to 1998-2000: 1995-97: OR=1.8 (1.3, 2.5) 1992-94: OR=1.8 (1.3-2.6) 1989-91: OR=1.09 (0.7, 1.7)	

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	CaPSURE (intermediate risk)	AS/WW							P=0.46 (Mantel-Haenszel test for trend)	A trend line for the frequency of undergoing WW is presented in the single Figure of the paper.
	CaPSURE (high risk)	AS/WW							P=0.14 (Mantel-Haenszel test for trend)	A trend line for the frequency of undergoing WW is presented in the single Figure of the paper.
Cooperberg[67] 2004 15169800	CaPSURE	WW ADT EBRT BT RP			13.8% 3.1% 16.1% 5.1% 63.8% (1989-92)	12.5% 12.8% 8.9% 8.9% 56.9% (1996-98)	7.9% 12.0% 6.8% 21.7% 51.6% (1999-2001)		P for trend <0.001 for all Tx except RP; P=0.0019 for RP.	Low-risk prostate cancer patients only (n=1990)
Cooperberg[90] 2008 18369637	CaPSURE	WW ADT EBRT BT Cryotherapy RP			4.4% 18.5% 21.6% 4.3% 6.2% 45% (1990-94)	3.9% 22.6% 14.6% 11.6% 4.1% 43.2% (1995-99)	2.7% 26.4% 12.3% 13.4% 4.8% 40.5% (2002-03)	3.0% 29.1% 10.9% 6.6% 8.5% 41.9% (2004-07)	P<0.001 (Cuzick test for trend)	High risk localized prostate cancer only
Cooperberg[69] 2010 20124165	CaPSURE (CAPRA 0-2)	AS/WW RP RT ADT			12.8% 61.6% 14.1% 4.7% (1990-94)	7.1% 61.6% 22.8% 6.3% (1995-99)	7.2% 59.6% 23.2% 7.1% (2002-03)	8.5% 59.5% 20.9% 6.4% (2004-07)	NR	

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
	CaPSURE (CAPRA 3-5)	AS/WW RP RT ADT			4.4% 54.9% 24.1% 9.4% (1990-94)	5.7% 47.2% 28.6% 14.7% (1995-99)	3.3% 48.6% 26.9% 17.2% (2002-03)	5.9% 48.7% 23.5% 14.9% (2004-07)	NR	
	CaPSURE (CAPRA 6-10)	AS/WW RP RT ADT			1.7% 27.5% 29.7% 36.7% (1990-94)	2.7% 22.1% 28.6% 41.9% (1995-99)	1.1% 23.3% 25.2% 43% (2002-03)	1.9% 22.9% 21% 45.5% (2004-07)	NR	
Greene[71] 2005 16194711	CaPSURE	RP Cryosurgery BT EBRT Orchiectomy LHRH agonist LHRH antagonist Antiandrogen 5α-RI WW				52% 2% 27% 13% 0% 3% 1% 1% <1% 1% (1997-99)	58% 3% 21% 10% <1% 4% <1% 1% <1% 2% (2000-03)		P=0.011 (chi-square comparing 2 periods)	
Mettlin[72] 1994 8062197	NCDB	TURP only RP RT RP+RT ADT RT+ADT RP+ADT RP+RT+ADT Other		32.1% 11.1% 26.8% 2.1% 17.0% 5.0% 0.7% 0.2% 5.0% (1985)	20.4% 23.7% 28.0% 2.4% 16.0% 3.4% 1.3% 0.2% 4.6% (1990)				NR	
Mettlin[73] 1996 8640686	NCDB	RP RT No Tx		12.4% 27.3% 38.0% (1987)	29.4% 30.6% 20.8% (1992)				NR	

Appendix Table C1.14. Treatment characteristics—changes in treatment patterns over time (continued)

Author Year PMID	Database	Treatment	1980-84	1985-89	1990-94	1995-99	2000-04	2005-10	Statistical analysis	Notes
Mettlin[74] 1995 8625214	NCDB	No Tx PT RT ADT Combo & other		39.1% 11.4% 27.6% 14.8% 7.2% (1987)	20.6% 29.1% 31.0% 11.4% 7.8% (1992)				NR	Total cases = 52596 in 1987
Mettlin[75] 1998 9781963	NCDB	RP EBRT BT ADT Other No treatment			31.6% 30.1% 1.4% 12.0% 4.9% 20.0% (1992)	34.1% 26.3% 2.2% 11.7% 4.1% 21.6% (1995)			NR	
Mettlin[91] 1999 10547563	NCDB	BT EBRT RP Hormone No treatment Other[b]			1.4% 32.5% 31.5% 16.8% 20.0% 8.5% (1992)	3.0% 26.3% 36.3% 22.8% 20.8% 6.9% (1996)			NR	All prostate cancer
Hamilton[77] 2011 20735387	POCS, NCI	BT (only) BT+EBRT EBRT (only) RP PADT WW				8.5% 6.4% 19.0 45.8% 7.6% 12.6% (1998)	12.3% 5.4% 20.1% 44.7% 8.5% 9.0% (2002)		NR	

Ordering of the studies follows Appendix Table C1.1.

5α-RI = 5α-reductase inhibitor; ADT = androgen deprivation therapy; AS = active surveillance; CAPRA = Cancer of the Prostate Risk Assessment; diff. = differentiated; OR = odds ratio; EM = expectant management; PMID = PubMed identification number; RP = radical prostatectomy; RT = radiation therapy; TURP = transurethral resection of the prostate; Tx = treatment; WW = watchful waiting.

[b] Includes subtotal prostatectomy, chemotherapy, immunotherapy, radioisotope, and unspecified treatments or combinations.

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Baylor College of Medicine and MSKCC, US[106] [15017211] 1984-2001	Prostate cancer diagnosed by needle biopsy or transurethral resection and Gleason sum 7 or less. All patients were eligible for definitive therapy in the form of RP or RT. No patient had significant comorbidities. The decision for deferred therapy was made by the patient and treating physician together based on the likely presence of small volume cancer.	Pathological features of the biopsy results, clinical stage and/or PSA influenced the decision to proceed with the deferred therapy protocol. Prospectively designed protocol of deferred therapy: office evaluations every 3 mo first yr and every 6 mo thereafter. It included digital DRE and PSA. Repeat TRUS guided sextant biopsy was recommended at 6 mo or if the patient showed DRE/TRUS or PSA abnormalities consistent with disease progression. PSA velocity was calculated from 3 separate recorded values in a 12-mo period.	Definitive treatment when objective progression or patients' requests. Definitive treatment included RP and RT.	A point system for evaluating progression, including Gleason score increase, PSA velocity, DRE/TRUS, and biopsy specimen.
BCCA, Canada[157] [9445192] NR	Patients who were placed onto a watchful waiting program. Patient who had received treatment (either hormones or PT) prior to the referral were excluded.	No fixed follow-up schedule; patients generally were seen every 3-6 mo as needed. PSA at diagnosis and all subsequent followup PSA were recorded.	NR	Clinical progression: an increase in palpable disease or T classification. Biochemical progression: PSA DT calculated by 2 methods.
Cleveland clinic, US[119] [21256549] 2004-2009	Low-risk features by D'Amico criteria; a repeat (confirmation) prostate biopsy of ⩾10 cores; favorable clinical and pathologic features at the diagnostic and repeat biopsy; absence of primary or secondary Gleason scores 4 or 5.	PSA every 6-12 mo, surveillance biopsy was usually performed every 2 yr or sooner.	Intervention was recommended to patients considering multiple parameters (PSA and PSA kinetics, changes in DRE, quantity of cancer in biopsy specimens, and biopsy Gleason score)	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Dana-Farber Cancer Institute, US[117] [21167525] 2000-2010	Clinically localized disease (T1c-T2c), Gleason score 6 or less with no pattern 4, <3 cores positive for cancer and no more than 50% of cancer in any core. No age, PSA values or PSA density exclusion criteria was used.	PSA and DRE every 6 mo, and 20-core biopsy every 12 to 18 mo Protocol with cure intent.	Patients with progression were offered surgery or radiotherapy.	Progression criteria: 1) 3 or more positive cores, 2) increased grade (Gleason score 7 or greater) and/or 3) more than 50% of any core involved with cancer.
Erasmus Univ. hospital, Netherlands[145] [7544841] ≤1990; 1993-2006[150]	Histologically confirmed cancer; Metastatic disease was excluded by a normal chest x-ray and a normal bone scan. The decision not to treat was made by the urologist in discussion with the patient and his family, with respect to patient age, general health, clinical stage and patient preference. All patients had estimated survival >1 yr. Men on AS who were detected within the screening	Usually followed clinically twice yearly (mean 2.7 annual visits, range 1.4 to 4.3) for physical exam including DRE and serum PSA and alkaline phosphatase levels. Bone scan and chest x-ray were repeated regularly and when clinically indicated. Follow-up regimens varied among local practices, data for this study were collected from chart reviews of medical history, DRE, dissemination studies, and PSA tests.[150]	Subjective progression, like obstructive micturition or pain, was considered for treatment decisions.[145] Note: The authors reported that of 13 patients with progression, 6 started treatment (5 for subjective symptoms; 1 for objective progression only). The authors also reported that PSA progression may serve as a trigger point to treatment.[150]	Local progression: symptomatic, increase in T category, increase in prostate size on DRE by 25%, or increase in ultrasound measured volume >40%. Metastatic progression: new bone lesion.

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
	program of the ERSPC. All men retrospectively met the following criteria: clinical stage T1c or T2, PSA ⩽15 ng/mL, and Gleason score <8. The choice of initiating and continuing an AS was patient desire and/or physician advice.[150] ERSPC-screening protocol: Men aged 50–75 yr with PSA measurements (threshold 3.0 or 4.0 ng/ml), and/or TRUS, and/or DRE, at 2- or 4-yr intervals. Abnormal findings lead to sextant prostate biopsies, the Finnish centers have changed to 10 or 12 biopsy cores. Prostatic volume is			

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
	measured by planimetric calculation during TRUS. After a PCA diagnosis, men are referred to the regular medical circuit (which may be the ERSPC centre), where decisions on treatment are made.[151]			
Four tertiary care academic medical canters,[m] US[109] [19233410] 1991-2007	Patients who would otherwise be considered for surgery or radiation due to a life expectancy >10 yr, and were defined as age ⩽75 yr, clinical stage T1-T2a, PSA ⩽10 ng/mL, ⩽3 positive cores at diagnostic biopsy, Gleason score ⩽6, no active treatment for a minimum of 6 mo after the second biopsy.	Office visits, review of general health and urinary symptoms, DRE and PSA every 6 to 12 mo, rebiopsies within 18 mo of starting AS and subsequently every 1 to 3 yr or prompted by a change in clinical status (e.g., significant and sustained PSA increase). MRI of the prostate was selectively used at diagnosis and every 1 to 3 yr after starting AS.	Criteria for recommending treatment were nonstandardized and physician specific.	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Freeman hospital, UK[141] [3191340] 1978-1985	Patients without symptoms after initial outflow tract surgery or biopsy.	Disease progression was monitored by history, physical exam, TRUS for T staging and prostate volume since 1983, serum acid and alkaline phosphatase and 6-monthly isotope bone scans.	No treatment until symptomatic progression.	NR
Hospitals in Manchester, UK[148] [11711356] NR	"Localized" (bone scan-negative) prostate cancer patients treated by watchful-waiting. All patients had PSA level < 50 ng/mL	Patients were followed-up at 6-month intervals. All patients underwent "multiple bone scans" (all negative), and hormonal manipulation was demanded by the protocol when the PSA rose to 50 ng/mL. DRE not always performed in patient with long-standing, stable PSA values.	Hormonal manipulation was demanded by the protocol when the PSA rose to 50 ng/mL.	Bone scan for metastases; PSA levels.
Howard University College of Medicine, US[138] [1600492] 1967-1989	Stage A and B prostate cancer patients who were in a prospective expectant management program.	3-mo intervals for the first 5 yr, then at 4 to 6-mo intervals thereafter. Each visit assessment included DRE, pap and since 1985 a PSA was done. Bone scans were done initially and annually thereafter. CT of the pelvis was used infrequently, primarily in patients who elected not to have any form of surgical therapy.	Management plan of watchful waiting for most patients until signs and/or symptoms of disease activity occurred. Any progressive changes in enzymatic activity and/or signs or symptoms of progression or metastasis (back pains or weight loss), or changes in rectal findings either by DRE or TRUS were treated despite evidence of a positive or progression of the bone scan.	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Johns Hopkins, US[113] [20439642] 1994-2008	PSA density ⩽0.15 ng/mL/cm3; T1C; 12-core biopsy: Gleason ⩽6; no Gleason pattern ⩾4; ⩽2 cores cancer positive; ⩽50% cancer in any single core (also included some men who did not meet these criteria due to personal preference or comorbidity)	Semiannual PSA and DRE; annual extended 12-core biopsy	Annual surveillance biopsy: Gleason ⩾7; or Gleason pattern 4 or 5; or >2 cores cancer positive; or single core >50% cancer. Patient request or encouraged to seek curative treatment if perineural invasion on biopsy.[137] PSA kinetics not used as a trigger for intervention.	Progression = unfavorable biopsy[249]

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Kagawa Medical Univ., Japan[153] [10765093] 1990-1998	Japanese patients with nonpalpable prostate cancer, detected by elevated PSA. Diagnosed histopathologically by TRUS-guided six sextant biopsy: (1) Gleason score ≤6; (2) 1-2 positive cores per 6 sextant cores; and (3) ≤50% involvement of any positive core	PSA doubling time based on 1st PSA >1 mo after biopsy. ≥3 values at intervals ≥1 mo apart for >6 mo. Exponential slope fitted by regression.	NR	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Kagawa Medical Univ., Japan[118] [18272471] 2002-2003	50-80 yr, initial PSA ≤20 ng/mL, 1-2 positive cores per 6-12 systematic biopsy cores, Gleason score ≤6, ≤50% cancer involvement in any core. Excluded if comorbidities: past stroke, unstable angina, DM uncontrollable with insulin, severe HTN, MI w/in 6 mo.	PSA every 2 mo for 6 mo, every 3 mo thereafter. Re-biopsy at 1 yr (no data beyond 1 yr)	PSADT ≤2 yr after 6 mo (based on all PSA or most recent 1 yr) Re-biopsy did not fit initial pathology criteria	NR
Kansas City VA, US[158] [21172105]] 2004-2009	Low-risk prostate cancer patients: stage T2 or less, Gleason ≤6, PSA <20 ng/mL, and percent of total tissue on biopsy positive for cancer <20%	PSA every 3 mo and a repeat TRUS guided prostate biopsy at 1 yr. All biopsies were performed using a standard 12-core biopsy scheme, however, an increased number of biopsies were taken for larger glands.	NR	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Kitasato Univ. Hospital, Japan[154] [11851612] 1991-2000	Biopsy-confirmed clinically localized prostate cancer	"a DRE," generally seen every 3-6 mo "as clinical circumstances dictated." Bone scintigraphy annually.	NR	Increase in T category, ⩾25% increase in prostate size on DRE, TRUS-measured volume increase >50%, positive finding on a bone scan, a blastic lesion seen on skeletal radiograph or soft-tissue metastasis by biopsy. Not biochemical progression (though PSA DT calculated).

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
McGill Univ., Canada[107] [18484590] 1987-2002	Patients with prostate adenocarcinoma with "favorable pathologic and biochemical parameters" ("clinically localized caner"[26]) or patients who decline definitive treatment. The reasons for AS included patient choice, limited life expectancy because of advanced age or poor medical condition, and presumed insignificant prostate cancer.	PSA and DRE was done every 3-6 mo. TRUS guided biopsy was done annually or when there was a change in DRE or PSA.	The decision to treat was attributed to local pathologic disease progression and patient preference. The decision to undergo treatment was based on the suggestion of disease progression because of a rising PSA level or clinical progression on DRE or repeated sextant biopsy.[26]	Progression of T stage to T2b or more, progression shown in biopsy: 3 positive cores or more, >50% cancer in at least 1 core, or Gleason pattern of 4 Development of metastatic disease.[26]

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Memorial Sloan-Kettering Cancer Center, US[115] [21167529] 1997-2009	Low-risk prostate cancer patients who were eligible for AS; PSA <10 ng/mL, no prostate biopsy Gleason grade 4 or 5, clinical state T1-T2a, ≤3 positive biopsy cores (minimum 10), no biopsy core containing >50% cancer involvement and confirmatory biopsy to reassess eligibility before starting AS	Semiannually with DRE, free and total PSA measurements, and a review of general health and urinary symptoms. Biopsy was routinely recommended within 12 to 18 months of starting AS and subsequently repeated every 2 to 3 yr or as needed.	Treatment was recommended when the patient no longer met study eligibility criteria during followup.	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Northern Stockholm, Sweden[140] [17467883] 1978-1982	Patients with clinically localized prostate cancer, diagnosed by biopsies and cytological assessment, initially managed with WW. Patients with palpable tumors (71% T1-2 and 29% with T3) were included in a prospective surveillance protocol with close follow-up. Bone scan and PAP were normal in all patients.	Followup was performed every 3 to 6 mo for the first 2 yr and every 6 to 12 months thereafter with DRE and PAP. Re-biopsies were done every year during the first 4 yr, and a bone scan was repeated every 12 to 18 mo.	Treatment was offered to the patients if clinical progression with symptoms occurred.	Clinical progression: positive bone scan or plain x-ray for the diagnosis of skeletal metastases.

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Orebro Medical Center, Sweden[139] [7933233] 1977-1984	Patients were given no initial treatment if the tumor was localized to the prostate gland without penetration of the capsule on DRE (stages T0 to T2) and there was no evidence of distant metastases. However, several restrictions were applied to those with a palpable tumor (stages T1 to T2). From March 1978 to Feb. 1979, patients >75 yr were not given any initial treatment (deferred treatment).	Every 6 to 12 mo, patients were followed with clinical exam, lab tests including PAP and bone scans. PSA was only performed during the last few yr.	Patients were treated hormonally if disease progressed for they had symptoms of progression.	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
PASS, US[121] [19758683] 2008 - ongoing	Clinical stage T1-2, NX/0, MX/0, no previous treatment for prostate cancer (including hormone therapy), ECOG performance status 0 or 1, elected AS as preferred management plan, If diagnosis ≤1 year of entry, at leat 1 biopsy with ≥10 cores, if diagnosis >1 year before entry, minimum of 2 biopsies, 1 ≤2 years before entry, no other malignancies except skin cancer or superficial bladder cancer	PSA every three mo; DER every 6 mo Rebiopsy at the baseline visit if a biopsy with >10 cores is not available, at 6-12 mo from the baseline visit, at 2 years, then every 2 years.	Active treatment will be offered to a participant who showed disease progression, but the participant may opt to remain on AS. If this occurs, a new PSA/stage/grade status will be assigned and further progression events will be determined using the new baseline criteria.	PSA DT < 3 year, any increase in Gleason grade, and clinical progression

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
PIVOT, US [146] [18783735] 1994-2002	Biopsy proven T1-T2/Nx/M0 prostate cancer of any histological grade, diagnosed within 12 mo, PSA≤50 ng/mL, ≤75 yr, bone scan negative for metastatic disease, estimated life expectancy >10 yr, medically and surgically fit for RP	Office visit & PSA every 6 mo Bone scan every 5 yr	Discouraged treatment for asymptomatic progression (e.g., per PSA)	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
PRIAS, Netherlands[120] [19817747] 2006 – ongoing	Originating from the ERSPC.[151] Histologically proven adenocarcinoma of the prostate; fit for curative treatment; PSA-level at diagnosis ⩽ 10 ng/mL; PSA density ⩽ 0.2 ng/ml/ml; clinical stage T1c or T2; adequate biopsy sampling according to biopsy protocol; Gleason score ⩽3+3=6; maximal 2 biopsy cores invaded with prostate cancer; willing to attend the follow-up visits.	PSA at 3 mo, DRE at 6 mo and standard rebiopsy after 1 yr.	PSA DT 0 to 3 yr, T state >2 or rebiopsy findings exceed study inclusion thresholds	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Princess Margaret hospital, Canada[156] [21211899] 1995-2010	PSA <10 ng/ml, clinical stage T1c-T2a, Gleason score <6, and ⩽3 positive biopsy cores (<50% of a core involved at initial diagnostic biopsy)	PSA was measured every 3mo for 2 yr and every 6 mo in stable patients. DRE was performed every 6 mo. A confirmatory biopsy was typically performed 12 mo after the initial biopsy and then every 2–3 yr until the patient reached 80 yr of age or refused treatment. All biopsies were performed by one of three dedicated uroradiologists using a standardized approach that did not depend on prostate volume Fist-time biopsies consisted of 6 cores before 2001 and 11 cores after 2001. Repeat biopsies consisted of 10 cores before 2001 and 15-16 cores after 2001.	NR	Pathologic progression: increased grade, increased number of cores to more than 3 or any core involvement >50%
ProtecT, UK[116] [19603015] 2000-2008	Clinically localized prostate cancer. Patients agreed to participate in RCT and were allocated to active monitoring group, or refused to be randomly allocated to groups and chose to be managed by monitoring.	PSA every 3 mo in yr 1, and every 6 mo thereafter; referred to biopsy if a PSA ⩾3 ng/mL; rebiopsy was not routine	The aim of active monitoring is "to identify developing cancers early enough to allow treatment with surgery or radiotherapy"[n] "Test results were reviewed annually, and patient and clinician decided whether to continue with monitoring"[152] (implied using PSA level or change and/or rebiospy results as triggers).	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Royal Marsden Hospital, UK[112] [15839912] 1993-2002	AS: Fitness for RP, T1-2, N0/X, M0/X, PSA≤20 ng/mL, Gleason ≤7. "Favorable prognostic characteristics and according to patient preference." WW: localized prostate cancer (any T stage, N0/X, M0/X, any PSA, Gleason score ≤7). Unsuitable for RP typically because advanced age or comorbidities.	WW: PSA and DRE every 6 mo AS: PSA and DRE every 3-6 mo for 2 yr, then every 6 mo. Repeat Bx not routine. Repeat imaging only if clinically indicated.	WW: Symptomatic prostate cancer progression AS: Rate of rise of PSA, according to judgment of each patient and clinician.	NR
Royal Mardsen Hospital, UK[124] [17850368] ≥2002	T1/2a, N0/X, Mo/X, PSA<15 ng/mL, Gleason ≤7 (primary Gleason ≤3), cancer in ≤50% of biopsy cores. (Patients were 50-80 yr). Fit for RP, Elected AS for initial treatment	PSA monthly in yr 1, every 3 mo in yr 2, and every 6 mo thereafter. DRE every 3 mo for 2 yr. TRUS-guided octant biopsy at 18-24 mo	PSA DT<4 yr, histological progression (see Definition of progression), or patient preference, or PSA velocity >1 ng/mL/yr [128]	Gleason score >7, primary Gleason ≥4, (initial Gleason 3+3, upgraded to Gleason ≥3+4)[136] or ≥50% biopsy cores positive.
SPCG-4, Finland, Sweden, and Iceland[144]	Patients with newly diagnosed	Followup was done every 6 mo in the first 2 yr, then every 1 yr. Followup included: a clinical examination, measurement of hemoglobin, creatinine, PSA, and alkaline phosphatase	Adjuvant local or systemic treatment was not given. TURP was recommended	Local progression: a transcapsular tumor growth was

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
[12226148] 1989-1999	localized prostate cancer, < 75 yr, with life expectancy of >10 yr, T stage of T0d, T1 or T2, eligible for radical prostatectomy, well differentiated to moderately well differentiated tumor, negative bone scan, PSA level < 50 ng/mL. After 1994, men with T1c tumors — according to the revised 1987 International Union against Cancer classification — were also eligible. Men with a poorly differentiated tumor were not eligible. Patients whose condition was diagnosed with an extended biopsy protocol were accepted if <25% of the	levels. A bone scan and chest radiograph were obtained every 1 yr after start of the study. After 1996, chest x-ray films were obtained annually for the first 2 yr. Rebiopsy was not routinely undertaken.[152]	as a treatment for local progression.	palpable; symptoms of obstruction of the flow of urine that necessitated intervention, or both.

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
	tumor was Gleason grade 4 and <5% grade 5.			
Taichung Veterans hospital, Taiwan[143] [12854876] 1983-1996	Men undergoing TURP for clinically benign hyperplasia of prostate with stage T1a prostate cancer.	After the introduction of serum PSA in 1990, 3-6 monthly PSA and DRE were used to detect the disease progression.	No treatment until there was evidence of cancer progression.	Abnormal DRE and/or progressive elevation of PSA "proved" by transrectal needle biopsy, or appearance of metastatic disease.
Toronto-SRCC, Canada[114] [11395227] 1995-2002 as a phase II trial 2003-ongoing as an observational open prospective cohort	Histological diagnosis of adenocarcinoma of the prostate within 12 mo of enrollment; no previous treatment for prostate carcinoma; clinical stage T1b-T2b N0 M0 (1997 TNM classification); PSA ≤15ng/ml; Gleason score ≤7. Protocol changes in inclusion criteria and additional information regarding the original criteria, reported in	Every 3 mo for the first 2 yr and every 6 mo thereafter	Clinical, histological or PSA progression triggered the offer of treatment based on age, extent of disease and comorbidities. Specific treatment protocol was not reported. Protocol changes in PSA DT assessment, reported in Klotz 2010:[125] For the first 4 yr of the study, PSA DT <2y was used as a trigger. This criterion identified 10% of patients as high-risk and was considered overly stringent. In 1999 the cut-off was increased to 3 yr. Protocol changes in PSA DT calculation, reported in Loblaw 2010:[130] From 1995 to 2002 PSA DT was calculated by a statistician using linear regression of all PSA values after the patient left the clinic and the 95% upper bound confidence limit of PSA DT	Clinical progression = at least one of the following: >2 times of the product of the maximum perpendicular diameters of the primary lesion as measured digitally; symptoms requiring TURP; development of ureteric obstruction; radiological or clinical evidence of distant metastasis. Histological progression = Gleason score upgraded to 8 or greater in the rebiopsy of the prostate at 18 mo post enrollment. PSA progression = when all the following were

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
	Klotz 2010:[125] Between 1995 and 1999, study was offered to low-risk patients (Gleason ⩽6 or less, PSA ⩽10 ng/ml) and to patients older than 70 yr old with PSA <15 ng/ml or Gleason ⩽3+4. Since January 2000, the study was restricted to low-risk patients only.		had to be <3 yr. Later PSA DT was calculated by physicians who used PSA fluctuations to determine whether PSA DT was "truly" <3 yr. In 2005 the group developed a general linear mixed model as a clinical decision making aid.[o]	satisfied: PSA DT <2 yr, based on at least 3 separate measurements over a minimum of 6 mo; final PSA >8 ng/ml; p-value <0.05 from regression of ln(PSA) on time. Additional information on biopsy frequency during followup, reported in Klotz 2010[125] and Krakowsky 2010:[129] Subsequent biopsies were performed every 3-4 yr to identify biologic progression. Patients with borderline PSA DT underwent biopsies more frequently. Between 1995 and 2000 sextant biopsies were used; since 2000, 10 to 14-core biopsies were performed using the Vienna nomogram.

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
UCSF, US[111] [18433013] >1991	Prostate cancer diagnosis, no prior therapy at another institution, primary therapy AS or no primary therapy (surgery, radiation, brachytherapy, androgen ablation) within 6 mo of diagnosis Patients selectively were offered AS if they met the following diagnostic criteria: PSA <10 ng/mL, Gleason sum ⩽6, absence of Gleason grade 4 or 5, cancer involvement of <33% of biopsy cores, and clinical T1/T2a tumor	Office visit w/DRE every 3 mo, PSA every 3 mo (usually), TRUS every 6-12 mo. ⩾2003: prostate biopsy every 12-24 mo ⩾2002: "regular" nurse practitioner contact to ensure surveillance compliance and address concerns and anxiety	Implied that there was not a specific protocol for intervention; active treatment based on disease progression	Increase in Gleason or PSA velocity >0.75 ng/mL/yr (also analyzed PSA velocity >2 ng/mL/yr and PSA DT<2 yr. Ultrasonography not used (too much inter-observer variability in lesion size) Gleason upgrade to ⩾4 (if (⩽6 at diagnosis) or ⩾4+3 (if 3+4 at diagnosis); PSADT ⩽2 or 3 yr[131] Gleason ⩾7 or ⩾33% of cores or >50% of any core[132]

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Univ. of Connecticut Health Center, US[108] [18707696] 1990-2006	Patients who elected WW or AS program. Men on WW were generally older with localized prostate cancer who did not desire aggressive intervention. Men on AS were generally younger with low-risk disease.	WW: no additional information. Patients on AS were followed with PSA on an average of every 6 mo. If PSA trending upward, the checks increased to every 3 mo depending on initial presentation and PSA trend. Rebiopsies recommended 2 yr after initial biopsy or if an increased in PSA >0.75 ng/dl, a change in DRE or at patient request.	Increase in tumor volume (increased number or percent of cores positive), progression in Gleason score, onset of urinary symptoms, change in DRE or patient request for definitive treatment due to anxiety related to increasing PSA trend.	NR
Univ. of Florida, US[149] [18263992] 2003-2006	Low-stage, low-grade disease (minimal disease on biopsy), severe medical condition with a life expectancy of <10 yr, and patient's desire.	Patients are followed every 3 mo with PSA and DRE annually. Repeat biopsy about 6 mo after the initial diagnosis.	Cancer progresses or symptoms become imminent.	NR
Univ. of Miami, US[110] [17850361] 1991-2007	Patients with clinically localized prostate cancer who elected for watchful waiting and to be treated only when disease progressed.[127] No strict guidelines for	DRE and PSA every 3-4 mo for 2 yr and every 6 mo subsequently. After 2000, a laterally directed and peripherally targeted TRUS biopsy of 10-12 cores was performed 9-12 mo after the first rebiopsy, and then every yr or earlier if there was a dramatic rise in PSA or a change on DRE.	Disease progression[110] Treatment is encouraged at an increase in tumor volume, Gleason score $\geqslant$7, or the presence of >2 positive cores.	Local stage progression detected by DRE and/or biochemical progression (PSA increase 25-50 %/yr) or systemic progression when metastases detected.[110]

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
	accruing patients on the AS protocol in the early yr. Generally, patients with a Gleason score ⩽7 and stage ⩽T2b were offered AS. Over the yr the inclusion criteria became narrower, i.e. Gleason score ⩽6, PSA ⩽15 ng/mL, stage ⩽T2 and low-volume disease (⩽50% of two biopsy cores). Slightly changes in eligibility criteria, reported in Soloway 2010[123] and Gorin 2011:[122] AS is offered to prostate cancer patients with PSA ⩽10 ng/mL, Gleason score ⩽6, ⩽2 positive biopsy cores with tumor ⩽20% in each core,			

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
	stage ≤T2 and age ≤80 yr.			
Univ. of North Carolina, US[155] 1991-1996	Patients with stage T1c prostate cancer who chose to be on expectant management. All patients had DRE enzymatic PAP values obtained prior to or at least 4 wk after prostatic biopsy, radionuclide bone scan, and the absence of a visible lesion upon transrectal US that was proven malignant histologically.	PSA was monitored at 3 mo, then every 6 mo. Hematocrit and creatinine were measured every 6 mo.	NR	Development of palpable disease in DER, gross hematuria, urinary tract infection, bothersome symptoms due to bladder outlet obstruction, metastatic disease as shown in physical examination or radiographic examination, PSA level increase in 3 consecutive measurements and the total increase was > 5 ng/ml.

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Watchful Waiting Study, US[147] [14501381] 1998-2003	Age <85 yr, biopsy proven prostate cancer within 48 mo, PSA <50 ng/mL, have not received any therapy including surgery, radiation, hormone or chemotherapy, have not been diagnosed with metastatic disease, at least 3-yr life expectancy, no history of any type of malignancy within the past 5 yr with the exception of non-melanoma skin cancer, liver and kidney function within 1.5 x upper range of normal, not taking > 50 ug selenium/day as supplement, Gleason score <8.	PSA every 3 mo.	Developing progressive disease or electing to initiate cancer therapy	NR

Appendix Table C2.1. Eligibility criteria, follow-up protocols, triggers for intervention and definition of progression in cohorts of active surveillance/ watchful waiting/other observational management strategies (continued)

Center, Country [PMID] Enrollment year	Eligibility criteria	Followup or monitoring protocol	Triggers for intervention/ active therapy	Definition of progression
Western General Hospital, UK[142] [8343901] 1978-1990	Early cancer as either incidental (T0/stage A) or localized *T1/stage B1/B2), non-metastatic (M0) disease with normal serum PAP.	Every 3 mo for clinical assessment, routine blood tests and measurement of serum markers. Chest X-rays, skeletal X-rays and bone scans were performed every 6 mo. Urinary flow rates and residual volumes were assessed if outflow obstruction was suspected.	Progression of disease and/or development of symptoms.	Development of metastases (M1) or elevation of PAP to more than 2 u/l.

NR = not reported; DT = doubling time; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; TURP = transurethral resection of the prostate; yr = yr(s); wk = wk(s); mo = mo(s); SRCC = Sunnybrook Regional Cancer Center; BCCA = the British Columbia Cancer Agency; DRE = digital rectal examination; WW = watchful waiting; AS = active surveillance; EM = expectant management; PAP = prostate acid phosphatase; PSA = prostate specific antigen; TRUS = Tans-rectal ultrasound; CT = computerized tomography; PSA = prostate-specific antigen; TNM = tumor-node-metastasis system; SRCC = Sunnybrook Regional Cancer Center; BCCA = the British Columbia Cancer Agency; ED = erectile dysfunction; PRIAS = Prostate cancer Research International; ProtecT = Prostate testing for cancer and Treatment; SPCG-4 = Scandinavian Prostate Cancer Group Study Number 4; UCSF=University of California at San Francisco; European Randomized Study of Screening for Prostate Cancer = ERSPC; VA = Veterans Affairs; MSKCC = Memorial Sloan-Kettering Cancer Center; PASS = tthe Canary Prostate Active Surveillance Study; PMID = PubMed identification number.

[m] Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami

[n] Source: http://www.epi.bris.ac.uk/protect/

[o] The model generates 2 reclassification curves (high and low risk) which, when overlaid over PSA data of each patient, defines 3 risk zones of high, intermediate and low risk of reclassification. A patient with a PSA consistently in the high risk zone is recommended to undergo treatment.

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Physicians factors affecting offer	Crawford[193] 1997 9301699	Random phone survey of 1000 men from the 30,000 men in the prostate cancer support group (US TOO), of which 780 men responded; and survey of 200 urologists (details of sampling population of urologists not provided)	WW (not explicitly provided)	• 83% patients and 76% urologists preferred aggressive therapy • Treatment options for men with localized disease and few comorbidities, urologists on the average would prefer RP (67%), RT (29%), WW (4%) • Different perspectives on whether treatment options were discussed: 20% patients felt treatment options were not discussed while 1% urologists felt treatment options were not discussed	• 200/335 urologists and 780/1000 patients responded to the survey • Urologists in this survey were not necessarily the urologists who took care of the patients in the survey
Patient factors affecting acceptance	Berry[191] 2003 12856636	Content analysis of 13 men in focus groups and 31 men in individual unstructured interviews; men were within 6 mo dx of localized prostate cancer; sample of 44 obtained from 68 eligible men from 3 urology clinics and flyer and newspaper announcements	WW (not explicitly provided)	• 20/44 men who relied on influential others (an individual whose illness experience and/or story had explicit influence on the participant's treatment decision) to make a treatment decision, 1 broadened the horizon to consider WW, 1 moved away from considering WW	• Small sample size
Patient factors affecting acceptance	Chapple[189] 2002 12133062	Interview 50 men from UK with all stages of prostate cancer, 4 of whom chose WW; sample chosen to include men at different stages of dx; recruited through GPs, urologists, support groups and charities; although great effort to recruit minorities, few volunteered	WW = no active treatment	• Few men who chose WW had consulted the Internet, concerned about the side-effects and uncertain treatment outcomes, and found physicians who were supportive of their decision	• Men with all stages of disease in UK • Small sample size

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Patient factors affecting acceptance	Davison[183] 2009 19136342	Qualitative description of interviews of 25 of 45 eligible men with low risk prostate ca on AS; sample of English speaking men with low-risk prostate cancer currently on AS recruited from the Prostate Centre at Vancouver General Hospital and the British Columbia Cancer Agency (large urban tertiary referral centers for the province of British Columbia)	Implied; details not provided (patients from 2 large tertiary care centers that support AS)	• MD description of prostate ca affects patient perception of the seriousness of the condition and affects treatment choice • MD recommendation most influential on patient decision to select AS • Concerns about impotency and incontinence affects treatment choice	• Small sample size • Limited applicability
Patient factors affecting offer	Demark-Wahnefried[190] 1998 9669815	Survey of 231 men (50% Black) with prostate cancer in N. Carolina; stratified design to recruit 240 men evenly sampled with respect to race (50% black; 50% white) and stage (local, regional and distant disease) within the N. Carolina Central Cancer Registry; eligibility included ages 50-74 y; dx'd between 1994-95; reside in a region comprising 63 contiguous counties where >20% were black; dx'd at 1 of 16 hospitals or clinics with IRB approval; cancer registry included phone contact	WW (not explicitly provided)	• WW discussed ⩾ high school vs. <high school education 59.5% v. 43.7% (P<0.05) • MD recommendation most influential in treatment decision (57%) (no differences between Blacks and Whites (no numerical data); urban vs. rural (62.3% vs. 43.9%, P=0.004)) • Differences NS in WW options discussed between rural and urban residents (53.7% vs. 51.9%) • Differences NS in WW options discussed between Blacks and Whites (48.7% vs. 56.1%)	• No statistical adjustment

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Patient and physician factors affecting acceptance	Diefenbach[185] 2002 11828358	Survey of 654 men (77% RT; 17% RP; 6% WW) with early stage disease recruited by either a urologist or radiation oncologist; pts presented to Fox Chase Cancer Center or an affiliated hospital for an opinion regarding treatment options; eligibility criteria included early-stage disease; not yet decided on a treatment; free of substantial comorbidity, and ability to communicate in English	WW (not explicitly provided)	• Most influential in reaching a treatment decision: physician recommendation (51%), advice from family and friends (19%), information from books and journals (18%), Internet (7%), disease and treatment factors (3%) • Patients who chose RP over RT or WW perceived prostate cancer as a significantly more serious disease (P <0.001)	• Unclear if WW was actively offered by urologists or radiation oncologists to patients since only 6% opted for WW
What MD would offer based on expected life expectancy	Durham[184] 2003 12835804	Survey (of screening behaviors, with case vignettes, piloted for understandability and face validity on 10 GPs) of GPs (in New Zealand), equalized urban vs. rural. 201 urban, 180 rural GPs responded. Survey sent to a random national sample of 575 New Zealand GPs, stratified to include equal number of rural and urban GPs; 66.3% responded	WW: not defined (non-curative)	• For men with localized prostate cancer, GPs responded that If life expectancy <10 y, WW would be suggested treatment (45%), followed by hormone (23%), RT (13%), prostatectomy (8%), other combinations (6%) • If life expectancy >10 y, WW suggested 3%; prostatectomy 53%, other combination 17%, RT 14%, hormone 8%	• Survey of GPs given theoretical cases (vignettes) • No data urban vs. rural • Survey response rate 66%

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Offer of WW by MD	Fowler[188] 2000 10866869	Survey ("pretested") returned by 504 urologists and 559 radiation oncologists; random sample of urologists (response rate 64%) and radiation oncologists (response rate 76%) in the AMA Registry of Physicians who practiced at least 20 h per week	WW = "expectant management"	• ~10-20% of urologists and radiation oncologists would recommend WW if PSA ~5 ng/mL and Gleason score 4 or 5 (Scenario was for a 65 yr man in good health, with negative DRE and no evidence of nonlocalized disease). • Almost no (0-1%) would recommend WW for those with higher PSA or Gleason scores. • No difference between urologists and radiation oncologists.	• Surveys sent to urologists and radiation oncologists were somewhat different • Survey response rate 64% (urologists) & 76% (radiation oncologists)
Offer of AS by MD; acceptance of AS by patient	Gorin[122] 2011 21215429	Survey of 185 men already on AS (unclear selection procedure) in a university-based urologic oncology practice (105 of 185, 57% responded); pts were asked whether the urologist who dx'd the cancer offered AS as a primary treatment alternative	DRE + PSA q 3-4 mo for the first 2 yr, then q 6 mo; annual bx; sooner if significant rise in PSA or change in DRE; treatment encouraged for ↑ tumor volume, Gleason ⩾7 or >2 positive cores	• AS offered by the MD who had made the initial dx in 38/105 (36%) • MD influence had the greatest impact on choosing AS (73%) • Concerns for incontinence (48%) and erectile dysfunction (44%) also reasons for choosing AS	• Non-validated survey instrument • Population already decided to enroll in AS • Had been on AS varying times (some >2 yrs) • Survey response rate 57%
Patient factors affecting acceptance	Holmboe[187] 2000 11089712	Open-ended interview of 102 men with localized disease who had made a treatment decision but had not yet received the treatment (88% RP, RT or ADT; 12% WW); sample obtained from 128 consecutive men with newly dx'd localized disease, pts drawn from a university, a VA, and 2 community urology practices	WW (not explicitly provided)	• 30% men stated that physician recommendation influenced their treatment decision • 59% of patients discussed WW (presumably with their physicians) • Fear of consequences most common reason (64%) for not selecting WW; some of the others were perceived elevated risk because of↑PSA or Gleason (12%); physician (12%) and/or family (4%) against WW	• Small sample size • Unclear details concerning WW

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Patient and physician factors affecting acceptance	O'Rourke[194] 1999 10370363	Qualitative description of interviews of 18 men with prostate cancer (dx'd within 6 wk; stage I or II; undecided choice of treatment) and their wives; they were referred by their urologists; sample recruited from 3 community practicing urology groups (urologists screened out cognitively impaired patients and spouses) in a western N. Carolina community; 19 couples were approached, 1 declined; "sampling aim not representative of the general population, but representative of the process of prostate cancer treatment selection"	WW (not explicitly defined)	• "The process of reaching a treatment decision was influenced by the urologists; second opinions [mostly concurrence between primary care physician and the urologist in this sample], and comparisons of self with others." • "Couples ruled out options based on formal and informal information, although sometimes inaccurate, personal and vicarious cancer experiences, and beliefs about cancer that were intricately tied to emotions and fears." • "Couples considered both their own individual histories and concerns and their shared life experiences." • "'Doing nothing' was ultimately rejected for the certainty they perceived to be associated with it: certain death, feared to be slow and painful."	• Small sample size

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Physician factors affecting offer	Ramsey[182] 2011 20959991	Survey of 238 men (multi-center) with newly dx'd localized T1-3 disease and 25 urologists concerning their office encounters (initial consultation vs. second opinion); survey of patients and their urologist in urology clinics in 3 states (Family And Cancer Therapy Selection study, Charleston, SC; Los Angeles, CA; San Antonio, Tx); pt recruitment occurred at the urology clinics; 423 invited, 240 met eligibility criteria; 238 analyzed (2 excluded for unspecified consultation type)	AS (not explicitly provided)	• Urologists recommended 0.52 more treatment options (SE 0.19, P <0.001) in initial consultation than in second opinion visit • For low risk disease, 25% urologists recommended AS, 77% recommended RP in initial consultation; 16% urologists recommended AS, 91% recommended RP in second opinion visit • Discrepancy between what physicians recommended and what patients heard physicians recommended: in patients for whom urologists recommended RP, 67% patients heard the recommendation; in patients for whom urologists recommended RT or ADT, ~25% patients heard the recommendation	• Cannot establish causality for more RP recommended by urologists; plausible that patients sought out urologists for a second opinion because the patients were more interested in RP • Applicability limited to patients/urologists in academic centers
Physician factors affecting offer	Steginga[192] 2002 11856106	Interview of 108 men with newly dx'd localized prostate cancer from 2 hospital clinics and 4 urology practices in Queensland, Australia; men were referred by their urologists to the project if they have localized disease (no metastatic disease on scans and X-rays) suitable for curative treatment; 131 consecutive men were referred; 119 eligible (newly dx'd localized disease, communicate in English; no head injury, dementia, concurrent psychiatric illness and cancer), 108 participated	WW (not explicitly provided)	• Unprompted recall of their urological consultation: 71% of the physicians discussed WW; 92% discussed RP, and 87% discussed RT	• Limited applicability to US

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Adherence to AS (actually receiving active treatment); clinical factors (perception of physician advice)	Zietman[186] 2001 11586206	Retrospective study of 198 men with early stage disease on WW in 2 institutions, 63 of whom ultimately received treatment; 53 of them responded to an 8-point phone questionnaire (10 did not respond because they had died, were too infirm or too elderly)	Surveillance/WW: no primary treatment with radiation, prostatectomy or androgen deprivation; DRE & PSA q4-6 mo (retrospective study)	• 81% believed that treatment was desired by the physicians, which was the primary cause of the change in plan. • In contrast, MD notes revealed that for only 24% was there documentation that MDs advocated therapy due to clinical or biochemical evidence of tumor progression. o 71% had PSA increase only and 11% had no progression evidence • Physicians more often perceived that treatment was initiated by patients (in abstract conclusions only)	• Nonvalidated telephone survey (not described) • Retrospective definition of WW • Only surveyed those who received therapy • Survey response rate 84% • Did not report on full survey results, including the intended purposes of influences that affected decision

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
Patient factors affecting treatment choice	Anandadas[195] 2011 21083643	Prospective, multicenter study of 768 men with low-risk T1/T2 cancer who were referred to centers that offered 4 treatment options (RP, RT, brachytherapy, or AS). Reasons for selecting treatment were recorded. 821 men from 7 different uro-oncology centers in UK were enrolled (all were initially referred for active treatment for early prostate cancer); 53 were recruited before brachytherapy was offered, they therefore were excluded from the analysis	”AS” (not explicitly provided)	• 61/768 (8%) chose AS • AS was more frequently chosen over time (2000: 0%; 2006 ~20%) • Reasons for choosing AS: o More convenient for lifestyle 16% (compared to all 4 treatments: 17%) o Fear of side effects 11% (compared to all 4 treatments: 9%) o Fear of other options 8% (compared to all 4 treatments: 12%) o Combination of reasons 16% (compared to all 4 treatments: 15%) o Other 23% (including “didn’t want active or invasive treatment”) (compared to all 4 treatments: 6%) o Unknown 25% (compared to all 4 treatments: 17%)	• Incomplete analysis of reasons for choosing AS. Most of those who selected “Other” chose AS, but the other reasons are unexplained. • It was not explicitly stated who provided reason for treatment choice, though implicitly the patient himself.
Patient factors affecting treatment choice	Xu[196] 2011 21830629	Patients with newly dx’d localized disease were recruited from 3 urology offices and 1 radiation oncology office in Detroit and also through flyers in hospitals and doctors’ offices; sampling aim was to maximize a broad range of experiences, no problem recruiting men who chose RP or RT, had difficulty recruiting men who chose WW; semistructured interview of 21 men with localized cancer <75 y; attempted to select a diverse sample of men based on race, age, social class, and income; easily recruited patients who chose	WW (not explicitly provided)	• “For most men, both black and white, treatment decision making occurred within an emotional context of fear and uncertainty and without systematic use of information.” • Reasons why WW chosen (n=2): o Maintain current QoL o Cancer small or slow-growing, may never cause problems; “If it’s not bothering me, why mess with it” o Will be able to act if cancer progresses o God will take care of it	• Only 2 patients with WW

Appendix Table C3.1 Studies on offer, acceptance, and adherence of active surveillance (continued)

Factors examined	Author Year PMID	Study approach	AS/WW definitions	Findings	Issues
		surgery or radiation, but only able to recruit 2 men who chose watchful waiting (1 of whom was diagnosed 2 years earlier)	o Doctor recommended it • Reasons why WW not chosen (n=19): o Only for older people with poor health; Too young; Self-perception of having good general health and the longevity of their relatives o Cancer will progress and kill if not actively treated; Too risky, like living with a "ticking time bomb"; Don't feel comfortable "doing nothing" with cancer; Would worry too much about cancer spreading o Doctor did not recommend it (only ½ remembered being told of WW) o Don't know about it o Friend(s) went through it had bad results o Family would be very upset o Avoid regret later • Some highly educated men had a harder time selecting a treatment because the information was "too general" and there was a "lack of consensus among experts"		

DRE = digital rectal examination; bx = biopsy; dx = diagnosis; GP = General Practitioner; PMID = PubMed identification number.

Appendix Table C3.2. KQ3 multivariable analyses

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Receipt of AS/WW versus alternative treatments								
Harlan[66] 2003 14532780	Clinical, social, insurance	CaPSURE	1989-2000	5365	Localized prostate ca	No active treatment	Logistic regression predicts WW vs. active treatment	1. low vs. high risk (D'Amico), OR=5.1 (CI 3.8, 6.9) 2. >75 yr vs. <65 yr, OR=14.3 (CI 9.1, 22.5) 3. comorbidity score >1 vs. 0-1, OR=1.43 (CI 1.1, 1.8) 4. private ins. vs. Medicare, OR=0.7 (CI 0.5, 1.0) NS: academic vs. community; black vs. white; education; income; in relationship
Meng[179] 2005 15821485	Clinical, insurance	CaPSURE	1989-2002	6074	Localized prostate ca – high risk	Not explicitly provided	Multinomial logistic regression WW vs. RP	1. ⩾70 yr vs. <70 yr, OR=49.4 (CI 13.2, 185.4) 2. PSA >20 vs. ⩽10, OR=4.6 (CI 1.7, 12.8) 3. Medicare + suppl vs. private, OR=9.2 (CI 2.1, 39.2) NS: Gleason, T stage, comorbidities, marital status

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Latini[170] 2006 16400651	Clinical	CaPSURE	1989-2004	5643	Biopsy-confirmed prostate cancer patients. Analysis of treatment choice was limited to men with localized disease (clinical stage T1 to T3a).	Not explicitly provided	Multinomial logistic regression RP vs. BT vs. EBRT vs. ADT vs. WW	No differences between Latino and non-Latino white men in primary treatment after adjusting for other variables (clinical risk, age, education, marital status, type of insurance, comorbidities, dx yr, and study site; P-value or estimates not reported). No other information reported for the association of ethnicity with WW as compared to other treatments.
Marr[165] 2006 16515991	Clinical	CaPSURE	1995-2003	5149	Men with localized prostate cancer (T3a or less with no evidence of lymph or distant metastases)	Not explicitly provided	Multinomial logistic regression predicting WW vs. RP	1. Heart disease vs. none OR=3.0 (2.2, 4.2) 2. Stroke vs. none OR=1.2 (0.7, 2.2) 3. Urinary conditions vs. none OR=1.4 (1.0, 2.1) 4. comorbidities: 1-2 other comorbidities vs. none OR=1.0 (0.7, 1.6) 3 other comorbidities vs. none OR=1.6 (0.9, 2.7) 6 or more other comorbidities vs. none OR=5.2 (1.8, 15.1) (results were not reported for other comorbidity groups) Estimates were adjusted for study site, dx yr, clinical risk, age, education, relationship status and BMI. Regression estimates or p-values were not provided for these variables.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Sadetsky[163] 2008 17893700	Delivery system	CaPSURE	1995-2006	2507	Newly diagnosed localized prostate cancer, >65 yr.	Not explicitly provided	Multinomial logistic regression EM vs. EBRT vs. ADT vs. BT vs. RP	Using RP as the baseline: Insurance status, insurance provider for predicting EM, - HMO vs. not, OR=0.62 (CI 0.29, 1.33) - PPO vs. not, OR=0.95 (CI 0.36, 2.51) - VA vs. not, OR=4.74 (CI 1.94, 11.55) - Medicate + supplement vs. not, OR=0.88 (CI 0.57, 1.37) - Medicare + FFS vs. not, OR=0.35 (CI 0.16, 0.78) - Medicare + HMO vs. not, OR=0.75 (CI 0.26, 2.13) - Medicare + PPO vs. not, OR=0.33 (CI 0.14, 0.77) Estimates were adjusted for education level, risk category, age at dx, income, relationship status, race/ethnicity, and yr of dx. No estimates or p-values were reported for these variables.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Dall'Era[160] 2009 19230923	Clinical, social	CaPSURE	1995-2007	5939	Patients with prostate cancer Patients undergoing cryotherapy were excluded.	Not explicitly provided	Binary logistic regression active treatment vs. WW/AS	Among patients with low risk: 1. social support, in permanent relationship vs. not, OR=1.82 (CI 1.13, 2.94) 2. insurance status, Medicare (with or without supplement) vs. private or VA, OR=0.49 (CI 0.34, 0.71) Overall cohort: Insurance status, Medicare vs. no Medicare, OR=0.53 (CI 0.35, 0.79) Multivariable models included: age at dx, race/ethnicity, education, relationship/marital status and insurance coverage. Results were only reported for relationship/marital status and insurance status; no estimates or p-values were reported for the other variables.
Moses[169] 2010 20100957	Clinical	CaPSURE	1995-July 2008	4284	Men with biopsy-proven prostate cancer, who reported a health-related quality of life questionnaire within 12 mo before selecting primary treatment by 2007	Not explicitly provided	Multinomial logistic regression with all variables significantly associated with receipt of treatment (AS vs. RP vs. RT, ADT vs. cryotherapy vs. TUMT) in a univariate test	AS vs. RP - White vs. African American: OR=0.52 (CI 0.22, 1.25); P=0.15 - Other vs. African American: OR=0.69 (CI 0.16, 2.97); P=0.62 - Other vs. White: OR=1.32 (CI 0.34, 4.64); P=0.15 Estimates were adjusted for risk (D'Amico level), age, health perception, number of comorbidities, education level, and type of insurance. Estimates or p-values were not reported for these variables.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Barocas[161] 2008 18707731	Clinical, social	CaPSURE	1999-2004	1421	Localized prostate cancer.	No treatment within 6 mo after dx	Binary logistic regression AS	1. age at dx, >74 yr vs. ⩽74 yr, OR=7.30 (CI 4.39, 12.21) 2. risk of disease (modified D'Amico), low vs. not low, OR=3.40 (CI 1.91, 6.04) 3. education level, high school or less vs. some or more college, OR=0.86 (CI 0.53, 1.41) "Low risk" = PSA<10ng/ml, stage T1 or T2a, PSA density <0.15, < 1/3 positive cores, and no Gleason pattern 4 and 5. The OR for patients who met all 4 criteria for low risk was 2.7 (CI 1.9, 3.8) vs. all other patients.
Konety[162] 2008 18343440	Clinical	CaPSURE	NR	11,261	Biopsy-proven prostate cancer	Not explicitly provided	Multinomial logistic regression WW vs. any other primary therapy	Model stratified for disease risk category (see paper for other risk categories): low risk patients, using WW as the baseline, - BT: ⩾75 yr vs. <75 yr, OR=0.234 (CI 0.161, 0.339) - BT + EBRT: ⩾75 yr vs. <75 yr, OR=0.109 (CI 0.025, 0.473) - EBRT: ⩾75 yr vs. <75 yr, OR=0.430 (CI 0.288, 0.641) - PADT: ⩾75 yr vs. <75 yr, OR=0.744 (CI 0.507, 1.090) - RP: ⩾75 yr vs. <75 yr, OR=0.014 (CI 0.008, 0.025) Results were adjusted for demographics and the number of comorbidities at dx. There was no significant interaction between age and comorbidity level.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
								Model stratified by number of comorbidities: no comorbidities, using WW as the baseline, - BT: ⩾75 yr vs. <75 yr, OR=0.165 (CI 0.068, 0.400) - BT + EBRT: ⩾75 yr vs. <75 yr, OR=0.139 (CI 0.038, 0.516) - EBRT: ⩾75 yr vs. <75 yr, OR=0.400 (CI 0.178, 0.898) - PADT: ⩾75 yr vs. <75 yr, OR=0.385 (CI 0.171, 0.866) - RP: ⩾75 yr vs. <75 yr, OR= 0.004 (CI 0.001, 0.015) See paper for other comorbidity categories Results were adjusted for demographic and risk covariates, and accruing site.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Roberts[180] 2011 21396507	Clinical, geographic	SEER-Medicare	2004-2005	8323	Men ⩾75 with localized disease, excluded HMO, dx from autopsy or death certificates, unknown poverty rate, Gleason score, PSA, or tumor stage	No RP, RT, or ADT	Unconditional multivariate logistic regression predicts active therapy vs. WW	1. Age, OR 0.92 (CI 0.91, 0.94) 2. Black vs. white, OR 0.57 (CI 0.47, 0.68) 3. Other than black vs. white, OR 1.34 (CI 1.07, 1.66) 4. Married vs. unmarried, OR 1.37 (CI 1.19, 1.58) 5. Unknown status vs. unmarried, OR 0.73 (0.61, 0.89) 6. NE vs. CA and SJ, OR 1.74 (CI 1.44, 2.09) 7. North Central vs. CA and SJ, OR 1.61 (CI 1.30, 1.99) 8. Other West vs. CA and SJ, OR 1.24 (CI 1.01, 1.53) 9. PSA 10-20 vs. >20, OR 1.21 (CI 1.00, 1.47) 10. Gleason 5-6 vs. 7-10, OR 0.40 (CI 0.36, 0.45) NS: poverty line, rural vs. urban, South, LA and SF, tumor stage, comorbidity
Shavers[166] 2004 15009794	Clinical, social, delivery system	SEER-Medicare	1994-1996	24,974	Black, Hispanic or White men with prostate cancer, ⩾65 yr, with continuous Medicare Part A & B coverage for ⩾1 yr prior to dx	No RP, RT, or ADT within 6 mo of dx	Binomial logistic regression predicts WW as initial therapy vs. all other treatments	1. Race/ethnic group, - Black vs. White, OR=1.3 (CI 1.1, 1.4) - Hispanic vs. White, OR=1.2 (CI 1.03, 1.4) 2. Stage, SEER historical stage, - in situ vs. local (1994), OR=8.8 (CI 3.5, 21.7) - regional vs. local (1994), OR=0.4 (CI 0.3, 0.4) - distant vs. local (1994), OR=0.2 (CI 0.1, 0.2) - local + regional (1995 to 1996) vs. local (1994), OR=0.9 (CI 0.8, 0.98) - unstaged/unknown, vs. local (1994), OR=1.2 (CI 1.1, 1.3)

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
								3. Grade, differentiation, - moderate vs. well differentiated, OR=0.3 (CI 0.2, 0.3) - poorly/undifferentiated vs. well differentiated, OR=0.1 (CI 0.1, 0.12) - unknown vs. well differentiated, OR=0.4 (CI 0.4, 0.5) 4. Life expectancy, <10 yr vs. ⩾10y, OR=1.4 (CI 1.3-1.6) 5. Age at dx, per yr, OR=1.1 (CI 1.07, 1.09) 6. Comorbidity, specific conditions, - CHF vs. not, OR=1.4 (CI 1.2, 1.6) - COPD vs. not, OR=1.4 (CI 1.2, 1.5) - dementia vs. not, OR=2.0 (CI 1.4, 3.0) 7. Mean inpatient comorbidity index, per unit, OR=1.9 (CI 1.5, 2.4) 8. Mean outpatient comorbidity index, per unit, OR=1.3 (CI 1.0, 1.6) 9. Marital status, - single vs. married, OR=1.5 (CI 1.4, 1.4) - 10. Income, median census tract income per yr, - <30,000 vs. ⩾40,000, OR=1.1 (CI 1.03, 1.2) - 30,000 to 39,000 vs. ⩾40,000, OR=1.1 (CI 1.03, 1.2) 11. Education, % of residents in census tract with less than high school education, - 20-29.99 vs. <20, OR=1.1 (CI 1.1, 1.2)

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
								- ⩾30 vs. <20, OR=1.2 (CI 1.1, 1.3)
Snyder[159] 2010 20734396	Clinical, social	SEER-Medicare	2000	13,769	Clinically localized prostate cancer, ⩾66 y, survived ⩾9 mo, on Medicare (not managed care)	No treatment within 9 mo of dx	Logistic regression predicts treatment vs. WW	Using WW as the reference treatment: compared to RP: 1. Age, per year, RR=0.73 (0.72, 0.75); P<0.001 2. Race, - black vs. white, RR=0.34 (CI 0.27, 0.44); P<0.001 - other vs. White, RR=1.52 (CI 1.13, 2.04) 3. Urban vs. rural, RR=1.54 (CI 1.20, 1.96) 4. SES highest vs. lowest quintile, RR=1.77 (CI 1.43, 2.19) 5. Grade poor vs. well differentiated, RR=13.38 (CI 9.26, 19.35) 6. Comorbidity, - 1 vs. 0, RR=0.84 (CI 0.71, 0.99) - ⩾2 vs. 0, RR=0.67 (CI 0.54, 0.83) Compared to RT: 1. Age, 2. Black vs. white, RR=0.39 (CI 0.31-0.50) 3. Urban vs. rural, RR=1.48 (CI 1.18-1.85) 4. SES highest vs. lowest quintile, RR=1.52 (CI 1.26-1.84) 5. Grade poor vs. well differentiated, RR=2.34 (CI 1.78-3.08) 6. Comorbidity 2+ vs. 0, RR=0.80 (CI 0.67-0.96) 7. Comorbidity, - 1 vs. 0, RR=1.11 (CI 0.96, 1.28)

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
								- ⩾2 vs. 0, RR=0.67 (CI 0.54, 0.83) compared to RT + ADT 1. Age, 0.90 (CI 0.90-0.91); P<0.001 2. Black vs. white, RR 0.39 (CI 0.31-0.50) 3. Urban vs. rural, RT RR 1.48 (CI 1.18-1.85) 4. SES highest vs lowest quintile, RT RR=1.52 (CI 1.26-1.84) 5. Grade poor vs. well differentiated, RT RR=2.34 (CI 1.78-3.08) 6. Comorbidity 2+ vs. 0, RR=0.80 (CI 0.67-0.96) 7. Comorbidity, - 1 vs. 0, RR=1.11 (CI 0.96, 1.28) - ⩾2 vs. 0, RR=0.67 (CI 0.54, 0.83) Additional information is provided in Table 2 of the paper for the comparison of ADT monotherapy vs. WW. All estimates were adjusted for SEER region.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Hamilton[77] 2011 20735387	Clinical; geographic	SEER-POC	2002	1139	Clinically localized prostate cancer	No therapy within 4 mo of dx	multivariate logistic regression predicts WW vs. any other treatment	1. Age ⩾75 vs. <60 OR=8.8 (CI 2.9, 26.76), P=0.008 (trend) 2. Not married vs. married OR=2.19 (1.03, 4.66), P=0.04 2. New Jersey vs. California OR=3.56 (CI 1.15, 11.03) 3. PSA ⩾20 vs. ⩽4.0 OR=0.18 (CI 0.04, 0.78), P=0.003 4. Gleason 8-10 vs. <6 OR=0.04 (CI 0.00, 0.32), P=0.03 5. Comorbidities ⩾1 vs. 0 OR=0.26 (CI 0.08, 0.89), P=0.03 NS: race

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Yan[177] 2000 10699903	Clinical	Survey of men diagnosed with prostate cancer (through the Washington U. PSA Prostate Cancer Screening Program) who chose 1 of 3 tx (RP, RT or WW)	1989-1998	1809 of 2345 provided followup questionnaire information	Screen-detected, clinically localized prostate cancer	Not explicitly provided	Multinomial logistic regression (WW vs. RP vs. RT)	1. Non-Black more likely (than Black) to choose RP than WW [OR=4.3 (1.7, 10.9)] or (nonsignificantly) RT than WW [OR=2.6 (0.86, 7.7)] 2. Clinical stage T2 more likely (than T1) to choose RP than WW [OR=3.0 (1.8, 4.8)] or RT than WW [OR=2.8 (1.6, 4.7)] 3. No urinary dysfunction more likely (than yes) to choose RP than WW [OR=1.8 (1.13, 2.8)] but NS RT vs. WW [OR=1.08 (0.66-1.8)] 4. No sexual dysfunction NS RP vs. WW [OR=0.83 (0.5, 1.3)] but less likely to choose RT than WW [OR=0.52 (0.30, 0.84)] 5. PSA level, for every 1 ng/mL increase (at dx) RP more likely than WW [OR=1.12 (1.04, 1.20)] and RT than WW [OR=1.15 (1.07, 1.23)] 6. Age, for every 5-yr increase RP less likely than WW [OR=0.21 (0.17, 0.27)] and RT less likely than WW [OR=0.49 (0.39, 0.63)] NS: marital status, education, income, indication for biopsy, and a Charlson-like comorbidity score.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Wolters[171] 2010 19739124	Clinical	Post hoc analysis of ERSPC	1993-2006	8010 (completed data set)	low, intermediate and high risk cancer	Not explicitly provided	Polytomous logistic regression predicts AS compared to RP	1.↑Age OR=1.19 (CI 1.17, 1.21) 2.↑PSA OR=0.30 (0.23, 0.39) 3. T2 vs. T1 =OR 0.33 (CI 0.28, 0.39) 4. PSA 50+ vs. ⩽4.0 OR=1.73 (CI 1.02, 2.94) 5. Gleason ⩾8 vs. ⩽6 OR=0.20 (CI 0.13, 0.32) NS: study arm; lymph node involvement
van den Bergh[181] 2009 19637245	Patient preference	Analysis of men in PRIAS	2007-2008	129	Men who decided on AS (PSA ⩽10 ng/mL, PSA density <0.2 ng/mL/mL, localized or nonpalpable disease, ⩽2 biopsy cores; ⩽3+3 Gleason)	Closely monitoring for disease progression to decide on initiating curative therapy	Multivariable regression to predict decisional conflict	Patients who perceived that physician played the most important role in shared decision-making process also had more doubts (high decisional conflict) regarding the choice for AS. Involvement of physicians in the decision-making process was assessed by a non-validated instrument.
Sommers[178] 2008 18704993	Patient preference	Survey of 428 eligible men with newly dx'd localized cancer from 2 RT and 2 urology clinics in Boston	2004-2007	167 had eligible + analyzable data	T1, T2N0M0, not yet treated	Not explicitly provided	Logistic regression predicts choice of WW vs. other treatments or undecided	1. Desire to avoid side effects main predictor of choice of WW (logistic regression coefficients not provided, P<0.05) 2. "Current bowel problem" was also a predictor of choice of WW (logistic regression coefficients not provided, P<0.05)

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Adherence to AS/WW								
Carter[176] 2003 14581423	clinical	DOD CPDR database	1991-2002	313	⩽70 yr, Gleason ⩽6 (no pattern 4), ⩽3 positive cores, ⩽T2, PSA ⩽ 20 ng/mL	Not explicitly provided	Multivariable Cox proportional hazard predicts definitive 2º treatment	1. T2c vs. T1a/b HR 16.4 (CI 3.16, 85.16), P=0.0009 2. PSA doubling time 2-5 yr vs. <2 yr HR 0.32 (CI 0.20, 0.52), P<0.0001 Median f/u 3.8 yr NS: age; PSA at dx; Gleason, race; FH; comorbidities
Latini[164] 2007 17632144	Clinical, social, delivery system	CaPSURE	1997-2002	105	Patients with biopsy-proven localized prostate cancer, who elected AS.	No treatment for ⩾6 mo after dx	Cox proportional hazards regression time-to-active treatment/ AS interruption	PSA velocity, ng/ml/yr, - -0.51-0.50 vs. <0.51, HR=0.402 (CI 0.092, 1.754); P=0.23 - 0.51-1.50 vs. <-0.51, HR=1.518 (CI 0.425, 5.419); P=0.52 - ⩾1.51 vs. <-0.51, HR=3.181 (CI 1.122, 9.016); P=0.03 P=0.01 Cancer anxiety change rate, HR=1.019 (CI 1.004, 1.035); P=0.01 The following NS variables were also considered in the model (HR estimates not provided): relationship; clinical risk group, D'Amico classification; BMI ; race; education; number of comorbidities; insurance; age at dx; PSA velocity x CA change rate (interaction).

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Meng[167] 2003 14634396	Social, clinical, delivery system	CaPSURE	1989-2001	457	Men with localized prostate cancer who chose WW as the initial treatment within 9 mo of the dx, no active treatment within 6 months of initiating WW and >6 months of study followup	Not explicitly provided	Cox proportional hazards models with backward stepwise regression (stay criteria p<0.1) for active treatment (WW interruption)	Of the 457 men initially treated with WW, 188 (41%) received subsequent active treatment at a median of 1.7 yr. 1. Disease risk (D'Amico), - High vs. low risk of prostate cancer: HR=2.75 (CI 1.84, 4.12); P<.0001 - Intermediate vs. low risk of prostate cancer: HR=1.51 (CI 1.05, 2.07); P=.028 2. Age, - 65-74 vs. <65, HR=0.70 (CI 0.41, 1.18); P=0.18 - ⩾75 vs. <65. HR=0.57 (CI 0.33, 0.96); P=0.035 3. Education level, - not college graduate vs. college graduate, HR=0.66 (CI 0.46, 0.94); P=0.021 - unknown vs. college graduate, HR=0.68 (CI 0.42,1.10); P=0.11
Koppie[168] 2000 10840429	Clinical	CaPSURE	NR	329	Men with biopsy-confirmed prostate cancer who elected WW as their initial treatment.	No therapy within 9 mo of dx	Cox proportional hazards regression, including an analysis of time-dependent predictors time-to-active treatment/ WW interruption	Cox regression using only baseline variables: 1. age, - 65-74 yr vs. <65 yr, HR=0.374 (CI 0.179, 0.784); P=0.009 - ⩾65 yr vs. <65 yr, HR=0.336 (CI 0.166, 0.679); P=0.002 2. clinical T stage at dx, - T2 vs. T1, HR =1.833 (CI 1.123, 2.992); P=0.015 - T3-T4 vs. T1, HR=1.149 (CI 0.440, 3.002); P=0.777 3. PSA at dx, ng/ml - 4.1-10.0 vs. 0-4.0, HR=3.064 (CI 1.352, 6.944); P=0.007 - 10.1-20.0 vs. 0-4.0, HR=3.680

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
								(CI 1.544, 8.769); P=0.003 - ⩾20.1 vs. 0-4.0, HR=6.864 (CI 2.587, 18.202); P<0.001 4. Gleason score at dx, - 7 vs. 2-6, HR=1.082 (CI 0.570, 2.053); P=0.809 - 8-10 vs. 2-6, HR=1.179 (CI 0.395, 3.515); P=0.7681 5. Disease risk, - intermediate vs. low, HR=NR; P=NR - high vs. low, HR=NR; P=NR 6. Race/ethnicity, Black vs. White, HR=1.220 (CI 0.451, 3.302); P=0.695 Cox regression using time-dependent covariates: Change in serum PSA vs. baseline, per unit, HR=1.99 (CI 1.18, 3.35). Results from models using changes in stage or PSA before treatment as time-dependent covariates were not reported. Estimates were adjusted for age, race, PSA at dx, clinical T stage, total Gleason score.

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Wu[175] 2004 14767282	Clinical	DOD CPDR database	1990-2001	1158	No metastases	No active treatment within 9 mo of dx	Multivariable Cox proportional hazard predicts 2º treatment (RP, RT, ADT) vs. staying on WW	1. Age at dx, per yr, HR=0.96 (CI 0.95, 0.98), P<0.001 2. Log(PSA), per unit, HR=1.43 (CI, 1.28, 1.60); P<0.001 3. Clinical stage: - T2 vs. T1, HR=1.32 (CI 1.04, 1.66), P=0.021 - T3+T4 vs. T1, HR=1.62 (CI 0.99, 2.63) ; P=0.054 NS: highest Gleason sum ; FH; comorbidity; dead or alive
van As[172] 2008 18342430	Clinical	Royal Marsden Hospital, UK; pts dx'd in a number of centers	2002-2006	326 recruited	Localized disease	PSA q 1 mo yr1, q 3 mo yr 2, then q 6 mo; Bx at 18 mo – 2 yr Treat if PSA velocity >1 ng/mL/yr; Gleason ⩾ 4+3 or >50% positive cores	Multivariable Cox regression with respect to radical treatment for patients who elected AS	free/total PSA ratio (P<0.001) and T stage (P=0.006) were independent predictors of time to radical treatment in patients on AS NS: initial PSA; PSA density; Gleason; % positive core; Number of positive cores; prostate volume Median f/u of 22 mo

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
El-Geneidy[174] 2004 15008720	Clinical	Portland VA; all pts dx'd with prostate cancer in one period; 187 on WW	1993-2000	175 on WW with T1 or T2 were analyzed	T1-2 on WW	Not explicitly provided	Multivariable Cox regression predicts curative treatment	1. Age 66-74 vs. >75 yr HR=5.0 (CI 1.13, 22.17), P=0.034 2. PSA doubling time <3 yr vs. >10 yr HR=2.73 (CI 1.19, 6.24), P=0.018 3. 34%-50% vs. <34% positive biopsies HR=2.47 (CI 1.14, 5.35), P=0.022 NS (univariate) : PSA ⩾4 vs. <4; Gleason ⩾6 vs. <5; T2 vs. T1; ⩾34% positive bx vs. <34% ; PSA density Median f/u 3.3 yr (range 0.1-8.6 yr).

Appendix Table C3.2. KQ3 multivariable analyses (continued)

Author yr PMI	Factors	Data source	Duration	Analyzed sample	Population characteristics	WW/AS definitions	Methods	Results as described in paper
Chose AS vs. randomization to available treatments								
Mills[173] 2006 16774847	Patient preferences	comparison of 180 men who refused randomization but selected AS with 138 men randomized to AS (from ProtecT study)	2001-2004	318	not reported	regular PSA	Multivariable logistic regression predicting "selecting AS" vs. "randomized to AS"	1. SES and baseline anxiety associated with selecting treatment: per decrease in SES from I to V, OR=0.68 (CI 0.49, 0.96); P=0.03. 2. Baseline anxiety (per unit increase on HAD scale) OR=0.93 (CI 0.87, 0.99); P=0.04) adjusted for baseline score, study center, age (further adjustment for marital status, SES had little impact (data not shown)) Applicability limited to predominantly white married, middle class men 50-69 yrs healthy for clinic testing.

Estimates are provided with 95% confidence intervals and p-values when available.
AS = active surveillance; BMI = body mass index; CI = confidence interval; DOD CPDR = Department of Defense Center for Prostate Disease Research; dx = diagnosis; ERSPC = European Randomized Study of Screening for Prostate Cancer; FH = family history; HAD = Hospital Anxiety and Depression scale; NS = non-statistically significant; PCOS = Prostate Cancer Outcomes Study; POC = Patterns of Care study; PRIAS = Prostate Cancer Research International: Active Surveillance study; ProtecT = Prostate Testing for Cancer and Treatment study; PSA = prostate-specific antigen; TUMT = transurethral microwave thermotherapy of the prostate; UI = PubMed unique identifier; WW = watchful waiting; yr = year.

Appendix Table C3.3. Lin 2009,[197] systematic review of patient decision aids

Author Year [PMID]	Lin 2009[197] [19841280]
Design	A systematic review of patient decision aids for prostate cancer treatment
Population	Men with low-risk prostate cancer who had the option of RP, RT, or WW
Intervention (Exposure)	Various decision aids (written information package, consultation with nurse or urologist, generic video, interactive computer program/CD-ROM decision aid, personalized multidisciplinary consultation, either stand-alone or in combinations)
Results	MEDLINE, CINAHL, Web of Science, and Cochrane Library initial search yielded 219 articles. 13 (3 RCTs (1 poor and 2 good per Jadad rating) and 10 nonrandomized trials) met eligibility criteria (inception through 3/2009). Key findings 1. Majority of DAs were developed de novo 2. The participants in general found the DAs to be informative. 3. One RCT reported a decrease in anxiety in participants in the intervention arm (written information package with discussion, a list of questions they could ask their physician, and an audiotape of the medical consultation) versus written information alone.[205] 4. One RCT found that there was no difference in satisfaction with treatment choice between those who received individualized DAs and those using a generic DA.[204] 5. One RCT found that the men in the DA arm selected their physician's treatment choice less often than those who received usual care.[203] 6. The nonrandomized studies reported that DAs appeared to increase patients' knowledge concerning prostate cancer and its treatments.
Comments	As noted by the systematic review authors: few high quality trials, heterogeneous outcome measures, and the quality of the information provided in the DAs themselves were not assessed and therefore whether these DAs met the quality standards set by the International Patient Decision Aids Standards Collaboration could not be determined.[206]
AMSTAR items	
A priori design?	Y
Two independent reviewers?	Y
Comprehensive literature search?	Y
All publication types and languages included?	N
Included and excluded studies listed?	N
Study characteristics provided?	Y
Study quality assessment performed?	Y
Study quality appropriately used in analysis?	Y
Appropriate statistical synthesis?	Y
Publication bias assessed?	N
Conflicts of interest stated?	Y

DA = decision aid; N = no; PMID = PubMed identification numner; RCT = randomized controlled trial; UI = unique identifier; Y = yes.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Randomized studies – clinical outcomes							
Bill-Axelson[213] 2011 21542742 RCT	SPCG-4	WW vs. RP	12.8 yr	695	Newly diagnosed prostate cancer patients, younger than 75 years, with life expectancy >10 years, T0d-T2, WHO well-moderately differentiated tumors.	Age: 65 yr Mean PSA: 13 Grade: WW, WHO 1, 47.7%; WHO 2, 52.3%; unknown, 0%. RP, WHO 1, 48.4%; WHO 2, 51.3%; unknown, 0.3% Stage: WW, T1b, 14.4%; T1c, 10.9%; T2, 74.4%; unknown, 0.3%. RP, T1b, 9.5%; T1c, 12.4; T2, 77.8%; unknown, 0.3%	A
Johansson[214] 2009 18783877 RCT	Ancillary investigation from SPCG-4 trial	WW vs. RP	Mean follow-up 4.1 yr (range: 1-8 yr)	376 (326 responders)	Newly diagnosed prostate cancer patients, younger than 75 years, with life expectancy >10 years, T0d-T2, WHO well-moderately differentiated tumors, PSA<50 ng/ml, negative bone scan, health status that would permit RP, life expectancy of >10 yr.	Mean age (range): WW, 65 yr (51-74); RP, 64 yr (48-74); PSA: NR Grade: WW, WHO 1, 47%; WHO 2, 53%, unknown, 0%. RP, WHO 1, 46%; WHO 2, 53%, unknown, 1% Stage: WW, T0, 14%; T1, 24%, T2, 60%; unknown, 2%. RP, T0, 12%; T1, 21%; T2, 61%; unknown, 5%	B Swedish component of SPCG-4

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Fransson[225] 2009 18985545 RCT	Ancillary investigation from UMEA1 trial	WW vs. RT	Median followup (range): WW, 114 mo (95-196); RT, 118 mo (77-188)	72 (54 responders)	Localized prostate cancer patients, no previous treatment for prostate cancer, no comorbid condition that would result in expected survival less than that of the normal population of the same age	Median age (range): WW, 78 (65-88); RT, 77 (54-87) PSA: WW, ⩽10, 37%; 11-20, 7%; 21-50, 26%, >50, 15%; unknown, 15%. RT, ⩽10, 26%; 11-20, 26%; 21-50, 26%, >50, 4%; unknown, 19% Grade: WW, WHO 1, 78%; WHO 2, 22%. RT, WHO 1, 70%; WHO 2, 30%, unknown, 1% Stage: WW, T1a, 15%; T1b, 19%; T1c, 0%; T2, 78%. RT, T1a, 0%; T1b, 15%; T1c, 4%; T2, 81%.	C Only 72 out of original 166 patients in the trial were included in this study.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Randomized study – treatment costs							
Andersson[237] 2011 21265595 Substudy of RCT	SPCG-4	WW vs. RP	Median followup 11.8 yr for the WW and 12.2 yr for the RP	212	<75 yr, life expectancy >10 yr, T0d-T2 disease, WHO well/moderately differentiated, PSA <50 ng/ml, no evidence of skeletal metastases on bone scan; patients from the trial were included if they resided in the counties where the two centers that randomized most patients were located (Örebro and Uppsala)	Age: WW, 64.4 yr; RP, 64.7 yr PSA: WW, <4, 27.6%; 4-6.9, 13.3%; 7-10, 16.2%; 10.1-20, 27.6%; >20, 12.4%; unknown, 2.9%. RP, <4, 16.8%; 4-6.9, 15%; 7-10, 19.6%; 10.1-20, 29.9%; >20, 15.9%; unknown, 2.8%. Gleason score: WW, 2-4, 21.9%; 5-6, 48.6%; 7, 24.8%; 8-10, 2.9%; unknown, 1.9%. RP, 2-4, 20.6%; 5-6, 49.5%; 7, 23.4%; 8-10, 2.8%; unknown, 3.7%. Stage: WW: T1b,12.4%; T1c, 7.6%; T2, 80.0%; unknown, 0%. RP: T1b,10.3%; T1c, 8.4%; T2, 81.3%; unknown, 0%.	B

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Observational studies – clinical outcomes							
Stattin[223] 2010 20562373 Retrospective cohort	NPCRSFS	Surveillance (AS and WW, in aggregate) vs. RP vs. RT	Median followup = 8.2 yr	6849	⩽70 yr, clinically localized disease (T1/2), N0/x, M0/x, PSA<20 ng/mL; Gleason score ⩽7	Mean age: surveillance, 64.7 yr; RP, 61.2; RT, 63.4 yr Mean PSA: surveillance, 7.6; RP, 8.2; RT, 9.3 Gleason score: surveillance, 2-4 or WHO I/II, 95.4%; 7, 4.6%; RP, 2-4 or WHO I/II, 82.3%; 7, 17.7%; RT, 2-4 or WHO I/II, 80.4%; 7, 19.6% Stage: WW: T1a, 16.4%; T1b, 4.5%; T1c, 50.5%; T2, 28.6%. RP: T1a, 1.5%; T1b, 1.4%; T1c, 53.1%; T2, 44.0%. RT: T1a, 1.2%; T1b, 2.0%; T1c, 43.7%; T2, 53.1%.	B
Litwin[229] 2002 12115317 Retrospective cohort	CaPSURE	WW vs. RP vs. RT	1.5 yr	452	Treatment within the first 6 mo of diagnosis, had completed at least two health-related quality of life surveys during the study	Age: 65.5 yr ± 8.3 yr PSA: 10.1 ± 11.2 Gleason score: 5.9 ± 1.2 Stage: T1, 30%; T2, 66%; T3/4, 4%	B

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Schymura[230] 2010 20403178 Retrospective cohort	CDC-NPCR POCS	WW vs. RP vs. RT	5 yr	3328	Histologically confirmed prostate adenocarcinomas, localized stage (clinically inapparent tumor, or cT1c/cT2, N0/x, M0/x; or pT1/pT2 N0/x M0/x)	Age: Under 70 yr, 57% PSA: NR Tumor grade: NR T1 or T2: 100%	B

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Hadley[207] 2010 20944078 Retrospective cohort	SEER-Medicare	RP vs. conservative management	1995-2003 Survival observed for up to 12 years; median survival time from date of diagnosis to Dec 31, 2007 (censoring date) was 78 mo (IQR=48 mo)	14302 used in survival models [a sample of 17815 patients was used for PS and IV analyses; then, exclusion of patients from geographic areas with fewer than 50 patients during the observation period and use of a lagged value in the IV analysis resulted in the exclusion of an additional 3513 patients]	patients with newly diagnosed prostate cancer, aged <75 yr, T1/2 tumor stage, receiving RP or conservative management within 6 months of diagnosis. Patients were excluded if they had "unusual histology," cancer diagnosis was based on death certificate or autopsy, were not from a SEER registry, had missing data on the month of diagnosis or date of death, were aged ⩽65 yr and had no data on the previous year; had incomplete Medicare Part A and Part B data because of managed care enrollment; had part A enrollment for only 1 year before or after diagnosis; had distant stage disease or not clinical T1/2 disease; or had received treatment with chemotherapy, radiation therapy, or hormone therapy, but without surgery.	Age: 60-69 yr, 50.4%; 70-74 yr, 49.6% PSA: NR Tumor grade: Well-dif., 7.9%; moderately dif., 70.4%; poorly dif., 18.8%; unknown, 11.4% Stage: T1, 63.6%; T2, 36.4%	B

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Wong[208] 2006 17164454 Retrospective cohort	SEER-Medicare	Active treatment vs. observation (a secondary analysis comparing radiation Tx and RP, separately, with observation was also reported)	1991-1999 12 yr followup	44,630	Patients aged 65 to 80 yr, with incident prostate cancer, stage T1/2. Patients were excluded if diagnosis was made at autopsy or death or if they had Medicare entitlement based on end-stage renal disease; were enrolled in a managed care plan from 3 mo before diagnosis to 6 mo after diagnosis; those with T3/4 disease, poorly differentiated or anaplastic tumors or metastatic disease, unknown tumor size; current reason for Medicare entitlement listed as disability or Medicare status were excluded. Patients who received ADT alone were excluded.	Median age: observation, 72.9 yr [IQR=69-77 yr]; active treatment, 71.0 yr [IQR=68-74 yr] PSA: NR Observation: well-dif., 25.87%; moderately dif., 64.13%. Active treatment, well-dif., 14.29%; moderately dif., 85.71% Observation: ≤T2a, 55.03%; T2b/c, 44.97%; active treatment, ≤T2a, 37.92%; T2b/c, 62.08%	B

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Berge[231] 2007 17178188 Retrospective cohort	SEER-Medicare	WW vs. RP vs. RT vs. ADT[a]	1991-92 5 yr	12,711	cT1-2 or pT1-3, age ⩾65, continuously enrolled on Medicare for the entire study period. Excluded patients enrolled in an HMO, those with primary treatment discrepancy between SEER and Medicare data, those who died before the end of the study period, patients with metastatic disease, and those with tumors clinically classified as extension through capsule or with positive lymph nodes (only for the non-RP groups). Patients with pT3 tumors were included only in the surgical arm to account for the lack of pathologic staging in patients receiving other treatments.	Median age: WW, 77; RP, 70; RT, 74 PSA: NR Grade: WW, well dif., 40.2%; moderately dif., 40.8%; poorly dif., 10.3%; undiff, 0.6%; unknown, 8.2%. RP, well dif., 10.2%; moderately dif., 67.6%; poorly dif., 20.5%; undiff, 0.5%; unknown, 1.2%. RT, well dif., 20.5%; moderately dif., 58.1%; poorly dif., 15.7%; undiff, 0.6%; unknown, 5.0%. Stage: WW, In situ, 0.9%; T1/2, 99.1; pT3, 0%. RP, In situ, 0.1; T1/2, 54.6; pT3, 45.4%. RT, In situ, 0%; T1/2, 100%; pT3, 0%.	C
Elliott[218] 2007 17570425 Retrospective cohort	CaPSURE	(RP, RP+EBRT, cryotherapy, BT, BT+EBRT, EBRT, or hormones) vs. WW	Median 2.7 yr (range 3 days to 10.9 yr)	6597	Newly diagnosed with prostate cancer between 1995 and 2006 with complete diagnostic and treatment clinical data available, and without a history of urethral stricture	Age: <60 yr, 25%; 60-59, 40%; ⩾70 yr, 35% PSA: ⩽4, 14%; 4.1-10.0, 62%; 10.1-20.0, 16%; >20, 8% Gleason score: 2-6, 65%; 7, 26%; 8-10, 9% Stage: T1, 53%; T2, 45%; T3a, 2%	B

[a] We did not extract data from the group of patients receiving primary ADT as the only initial therapy.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Schapira[226] 2001 11242319 Prospective cohort	4 academically-affiliated Wisconsin hospitals, including 2 VA Medical Centers	RP vs. RT vs. expectant management	3 and 12 mo	113 (pre-treatment) 112 (3 mo) 102 (1 yr)	⩾40 years of age, newly clinically localized prostate cancer (AJCC stage I or II). Exclusion criteria: Unable to speak English, a clinical diagnosis of dementia, or unable to verbally communicate. Dropouts: 6 patients died before the end of the study due to complications from radiation proctitis and cystitis after prostate cancer treatment with external beam radiation (n=1), myocardial infarction (n=1), bladder cancer (n=1), and undetermined causes (n=3). Other reasons for dropping out included geographic relocation (n=4, development of a new an serious illness (n=3), progression of an underlying comorbidity (n=1), and lost to followup (n=7)	Age: 69 (45-85) yr Median PSA (IQR) – RP: 7.6 (4.9-11.1) RT: 7.1 (4.9-12.3) EM: 7.9 (3.2-10.1) Gleason score in RP, RT, and EM groups, respectively- 2-4: 30%, 16%, 23% 5-6: 49%, 51%, 54% 7: 19%, 29%, 8% 8-10: 3%, 4%, 15% TNM Stage: EM, T1, 55%; T2, 45%. RP, T1, 55%; T2, 45%. RT, T1, 43%; T2, 57%	C Selection bias: 19% eligible patients were not contacted for a variety of reasons; dropout rate 12%, 9%, 7% in RP, RT, and RM group, respectively

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Rice[232] 2011 21872499 Retrospective cohort	Center for Prostate Disease Research (CPDR) database	WW (with or without secondary treatment) vs. EBRT vs. RP	Median 6.4 yr (all groups) Mean: WW without secondary treatment, 5.3 yr; WW with secondary treatment, 8.4 yr; EBRT, 7.0 yr; RP. 7.2 yr (P<0.0001 between groups)	770	12,081 patient records from the CPDR databse of men diagnosed with prostate cancer from 1989 to 2009. Of these 3650 were found to be ⩾70 years old. Patient who met the D'Amico criteri for low-risk disease (stage T1-2a, Gleason score ⩽6, and PSA <10 ng/ml) and were managed with primary RP, EBRT, or WW were selected for analysis. WW was defined as patient who declared the intent to pursue such pathway and not undergone definitive treatment within 9 months of diagnosis; this group was further stratified into patients who subsequently received or did not receive secondary treatment. Exclusion: any of the risk stratification criteria were missing or <6 months followup since primary treatment.	Mean age at survey: WW without secondary treatment, 75.7 yr; WW with secondary treatment, 74.5 yr; EBRT, 74.1 yr; RP, 72.2 yr (P<0.001 between groups) PSA (ng/mL): WW without secondary treatment, 4.7; WW with secondary treatment, 5.6; EBRT, 6.0; RP, 5.3 (P<0.001 between groups) Gleason score: NR Stage (P=0.32 between groups): T1: WW without secondary treatment, 66%; WW with secondary treatment, 61%; EBRT, 60%; RP, 57% T2a: WW without secondary treatment, 34%; WW with secondary treatment, 39%; EBRT, 40%; RP, 43%	C Significant difference in followup time between groups; self-selection bias likely; secondary treatment in WW group not properly accounted for in analyses.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Thong[224] 2009 19747357 Retrospective matched cohort	Eindhoven Cancer Registry (ECR)	"AS"[b] (long-term survivors) vs. EBRT(long-term survivors)	Mean 8 yr	142	All eligible patients diagnosed with prostate cancer between 1994 and 1998 from ECR. Excluding persons who had died before Nov. 1, 2004. For the purpose of this study, a sample of patients who would be suitable for management with AS according to the following criteria were selected: stage ⩽2 and a tumor grade of ⩽2 as determined with a biopsy at diagnosis. These patients thereafter received either no active treatment or at most, a TURP after diagnosis were matched with patients who had received EBRT as a primary treatment at diagnosis on (a) cancer stage, (b) tumor grade, (c) age at diagnosis (within 2 yr).	Mean age at survey: "AS," 75.8 yr; RT, 75.9 yr PSA: NR Grade: "AS," TNM Grade 1, 80.3%; TNM Grade 2, 19.7%. RT, TNM Grade 1, 80.3%; TNM Grade 2, 19.7%. Stage: "AS," stage 1, 67.6%; stage 2, 32.4%. RT, stage 1, 69%; stage 2, 31%.	B Of 128 AS survivors, 71 returned survey (55%)

[b] Although the authors referred to this group as "active surveillance" the study did not report following a predefined monitoring protocol; furthermore, patients in this group "received either no active treatment or at most, a TURP after diagnosis." For these reasons we did not consider this a comparative study of AS.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Lubeck[227] 1999 9886609 Prospective cohort	CaPSURE	Observation (WW) vs. RT vs. RP vs. hormonal therapy[c]	2 yr	692 (513 of whom received RP, RT or WW)	Newly diagnosed prostate cancer at enrollment in CaPSURE, with a minimum of 2 completed HRQoL questionnaires since study entry.	Mean age: WW, 72.1 yr; RT, 70 yr; RP, 62 yr PSA: NR Grade: NR Stage: WW, T1N0M0, 26%; T2N0M0, 64%; T3/4N0M0, 6%; TanyN+M+, NR; unknown, 4%. RT, T1N0M0, 39%; T2N0M0, 55%; T3/4N0M0, 5%; TanyN+M+, 1%; unknown, NR. RP, T1N0M0, 27%; T2N0M0, 68%; T3/4N0M0, 1%; TanyN+M+, NR; unknown, 3%.	C Only controlled for age

[c] Information was not extracted for patients who received hormonal therapy as their primary treatment.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Bacon[228] 2001 11586228 Prospective cohort	Health Professionals Followup Study	WW vs. EBRT vs. BT vs. RP vs. hormonal therapy vs. other treatments[d]	Up to 3 years	146 (patients with longitudinal data included in multivariable analyses; of these 452 received RP or WW)	The study was based on the Health Professionals Followup Study (a cohort study of male dentists, veterinarians, pharmacists, optometrists, osteopaths, podiatrists) aged 40-75 yr. Patients with localized prostate cancer based on independent review of medical records and pathology reports. Longitudinal comparisons were reported only for the subgroup of patients participating on an earlier substudy (after 1995) who also provided information in 1998.	Mean age at post-treatment survey: WW, 75 yr; RP, 68 yr PSA: NR Grade: WW, GS2-4, 16%; GS5-6, 42%; GS7-10, 19%; unknown, 23%. RP, GS2-4, 7%; GS5-6, 57%; GS7-10, 29%; unknown, 7%. Stage: WW, T1, 0; T2, 74%; "unspecified," 6%; unknown, 19%. RP, T1, 0; T2, 95%; "unspecified," 0%; unknown, 5%. [Data were extracted for the overall study population; separate data were not reported for the 146 patients with available longitudinal data]	C Only controlled for age

[d] We only extracted information on the comparison of WW with RP because RP was used as the baseline treatment in multivariable analyses in this study (i.e. direct comparisons were only possible between RP and other treatments).

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Albertsen[209] 2007 17296379 Retrospective cohort	Connecticut Tumor Registry	WW vs. RP[e]	Median followup: 13 yr (IQR=12.8, 13.9)	1618 (of whom 114 received no initial therapy and 802 were intended to receive RP)	Connecticut residents ⩽75 yr, with clinically localized prostate cancer, and initial PSA<50 ng/ml, treated with WW, RT or RP.	Median age: WW, 70 yr; RP, 65 yr Median PSA: WW, 6.6; RP, 9.1 Gleason score: WW, GS2-4, 17%; GS5, 15%; GS6, 46%; GS7, 11%; GS8-10, 11%. RP, GS2-4, 3%; GS5, 5%; GS6, 49%; GS7, 29%; GS8-10, 14%. Stage: NR	B
Observational studies – treatment costs							
Snyder[159] 2010 20734396 Retrospective cohort	SEER-Medicare	WW vs. RT only vs. Hormonal only vs. RT+Hormonal vs. Surgery±other	5 yr	13,769 (with cancer) + 13,769 (control group)	Localized prostate cancer diagnosed in 2000, 1st or only cancer in registry, survived ⩾9 mo, Age ⩾66 yr, in Medicare (not managed care) (Matched controls w/o cancer)	Mean age: WW, 77 yr; RT 74 yr; hormonal Tx, 79 yr; RT + hormonal Tx, 74 yr; surgery, 71 yr PSA: NR Tumor grade: Well diff., 5%; moderately diff., 69%; poorly diff /undiff, 22%; unknown, 4% Stage: NR (clinically localized 100%)	C

[e] This study compared three treatment strategies: observation (WW), RT and RP. Information on the comparison of WW versus RT was included in the previous AHRQ report[8] discussed in the section "Findings from previous systematic reviews" of the main text of the present report. Here, we extracted information on the comparison of WW versus RP.

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Penson[235] 2001 11248628 Retrospective cohort	CaPSURE	WW vs. active treatments RP monotherapy, RP + neoadjuvant hormone therapy; RT monotherapy, RT + neoadjuvant hormonal therapy; medical pADT, orchiectomy, medical pADT followed by orchiectomy]	1 yr	235	Patients enrolled in CaPSURE at the time of diagnosis, with T1c or T2 tumors, and complete resource date during followup	Mean age: 69.0 PSA: <10, 59%; 10-20, 27%; >20, 12%; unknown, 2% Gleason score: 2-4, 7%; 5-6, 58%; 7-10, 32%; unknown, 3% Stage: T1c, 25%; T2a/b, 44%; T2c, 31%	C

Appendix Table C4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Study duration	Sample size (total)	Inclusion criteria	Population description: Age PSA (ng/mL) Tumor grade Stage	Quality Comments
Wilson[236] 2007 Retrospective cohort	CaPSURE	WW vs. RP vs. cryotherapy vs. BT vs. EBRT vs. ADT vs.	5.5 yr	4553	Newly diagnosed prostate cancer patients identified within 6 months of diagnosis. Additional inclusion criteria were not reported; however, from the descriptive statistics of the population it appears that at least some T3a patients were included (the proportion is not reported). For this reason, when possible, we extracted data (e.g., cost outcomes) only for the subgroups of patients classified as "low" or "intermediate" risk, because these subgroups do not include T3a tumors (based on the study definitions of risk categories).	Age: <55, 11%; 55-65, 33%; 65-75, 40%; >75, 16% PSA: NR Gleason score: NR Stage: NR	C

ADT = androgen deprivation therapy; AS = active surveillance; BT = brachytherapy; CPT-4 = current procedural terminology, 4th edition; Dec. = December; Dif. = differentiated; EBRT = external bean radiation therapy; HCPCS = Healthcare Common Procedure Coding System; ICD = International Classification of Diseases, 9th edition; IQR = interquartile range; IV = instrumental variable; mo = months; NPCRSFS =National Cancer Register of Sweden Follow-up Study; NR = not reported; POCS = Patterns of Care Study; PS = propensity score; PSA = prostate specific antigen; yr = year; RCT = randomized controlled trial; RP = radical prostatectomy; RT = radiation therapy; WW = watchful waiting.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Prostate-cancer specific mortality							
Bill-Axelson[213] 2011 21542742 RCT	SPCG-4	RP vs. WW	Death from prostate cancer	12.8	RP: 347 WW: 348	RR=0.62 (0.44, 0.87); P = 0.01	None (RCT)
Hadley[207] 2010 20944078 Retrospective cohort	SEER-Medicare	RP vs. conservative management	Death from prostate cancer from SEER records.	Survival was observed for up to 12 years; mean survival time free of all cause-death = 83.0 mo; median survival time from diagnosis to Dec 31st, 2008 (censoring date) = 78 mo (IQR = 48 mo)	RP: 11,936; conservative management: 5879 [calculated based on the proportion of patients treated with each modality, for the overall population]	Unweighted regression analysis: HR=0.62 (0.50, 0.79); P <0.001 PS reweighted analysis using IPTW: HR=0.63 (0.55, 0.71); P <0.001 PS reweighted analysis using SMRW: HR=0.72 (0.57, 0.91); P <0.001 IV regression using the previous year's local area treatment pattern for conservative management as an instrument: HR=1.37 (0.15, 12.5); P = 0.78)	PS: age, race/ethnicity, marital status, tumor characteristics, previous health problems (based on NCI combined comorbidity index and Medicare reimbursements in the 12 months before diagnosis), year of diagnosis. These variables were included in all multivariable models. Instrumental variable: the lagged (previous year's) local area treatment pattern for conservative management.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Stattin[223] 2010 20562373 Retrospective cohort	NPCRSFS	RP vs. surveillance	Death from prostate cancer as "underlying cause of death," data obtained from the Cause of Death Register or review of death certificates	Median followup 8.2 yr (IQR=7.1-9.7 yr)	Surveillance: 2021 RP: 3399	HR=0.49 (0.34, 0.71)	Age at diagnosis, comorbidity, socioeconomic group, risk group.
Albertsen[209] 2007 17296379 Retrospective cohort	Connecticut Tumor Registry	RP vs. no initial treatment	Prostate cancer death; dates and causes of death were obtained from the Connecticut Tumor Registry	Median followup 13.3 yr (IQR=12.8-13.9 yr)	WW: 114 RP: 802 intended (596 received)	HR=0.29 (0.17, 0.52)	Stratification by D'Amico risk category, Gleason score, PSA, clinical stage, age at diagnosis, and Charlson comorbidity score [Gleason score values were based on a contemporary (2003) re-grading of the pathology slides]
All-cause mortality							
Bill-Axelson [213] 2011 21542742 RCT	SPCG-4	RP vs. WW	Overall mortality	12.8	RP: 347 WW: 348	RR=0.75 (0.61, 0.92); P=0.007	None (RCT)
Schymura[230] 2010 20403178 Retrospective cohort	POCS CDC-NPCR	RP vs. WW	5-year survival	5 years	RP: 1321 WW: 619	HR=0.43 (0.32, 0.59)	Age at diagnosis, race/ethnicity, marital status, registry location, PSA value, Gleason score, comorbidity score

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Hadley[207] 2010 20944078 Retrospective cohort	SEER-Medicare	RP vs. conservative management	Death from any cause from Medicare claims.	Survival was observed for up to 12 years; mean survival free of cancer-specific death = 73.2 mo	RP: 11,936; conservative management: 5879 [calculated based on the proportion of patients treated with each modality, for the overall population]	Unweighted regression analysis: HR 0.68 (0.63, 0.74); P <0.001 PS reweighted analysis using IPTW: HR 0.65 (0.62 ,0.68); P <0.001 PS reweighted analysis using SMRW: HR 0.68 (0.63, 0.75); P <0.001 IV regression using the previous year's local area treatment pattern for conservative management as an instrument: HR=0.92 (0.39, 2.17); P=0.78	PS: age, race/ethnicity, marital status, tumor characteristics, previous health problems (based on NCI combined comorbidity index and Medicare reimbursements in the 12 months before diagnosis), year of diagnosis. These variables were included in all multivariable models. Instrumental variable: the lagged (previous year's) local area treatment pattern for conservative management.
Stattin[223] 2010 20562373 Retrospective cohort	NPCRSFS	RP vs. surveillance	Death from any cause, data obtained from the Cause of Death Register or review of death certificates	Median followup 8.2 yr (IQR=7.1-9.7 yr)	Surveillance: 2021 RP: 3399	HR=0.49 (0.41, 0.57)	Age at diagnosis, comorbidity, socioeconomic group, risk group.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Wong[208] 2006 17164454 Retrospective cohort	SEER-Medicare	RP vs. observation	Overall survival = interval from the date of diagnosis to the Medicare dare of death. Patients were censored at Dec. 20, 2002.	12 yr	RP: 13,292 Observation: 12,608	HR=0.50 (0.47, 0.53)	PS: age at diagnosis, SEER site, year of diagnosis, tumor size, tumor grade, marital status, residence in an urban setting, race, income, educational achievement, and 44 categorical variables encoding comorbidities. The authors reported a statistically significant interaction between tumor size and grade. For treatment subgroups (RP and radiation Tx) separate PS were built and used as covariates in the Cox regression models.
Rice[232] 2011 21872499 Retrospective cohort	CPDR database	WW vs. RP	NR	Mean: WW, 5.3 yr; yr; RP. 7.2 yr	WW without secondary treatment: 214 WW with secondary treatment: 110 RP: 194	Multivariable Cox proportional hazards model predicting overall mortality: WW without secondary Tx vs. RP (reference): HR=1.938 (1.185, 9.168), P=0.008 WW with secondary Tx vs. RP (reference): HR=0.807 (0.462, 1.407), P=0.45	Age at diagnosis, PSA at diagnosis, race/ethnicity, number of comorbidities, T stage, and treatment groups (RP, EBRT, WW with secondary Tx; WW without secondary Tx).

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Incidence of distant metastases							
Bill-Axelson[213] 2011 21542742 RCT	SPCG-4	RP vs. WW	Metastatic lesions that were visible on a bone scan or histologically confirmed soft-tissue metastases outside the pelvic area	12.8 yr	RP: 347 WW: 348	RR=0.59 (0.45, 0.79); P <0.001	None (RCT)
Morbidity of primary treatment							
Johansson[214] 2009 18783877 RCT	Ancillary investigation from SPCG-4 trial	RP vs. WW	Physical symptoms (erectile dysfunction, weak urinary stream, urinary leakage, fecal leakage)	Mean followup: 4.1 yr (range: 1-8)	RP: 189 (166 responders) WW: 187 (160 responders) [some of the responders did not reply to all questions]	Last 6 mo; 2-3 yr post randomization Erectile dysfunction: RR=2.2 (1.5, 3.2) Weak urinary stream: RR=0.6 (0.4, 1.2) Urinary incontinence: RR=3.7 (1.6, 8.5) Fecal leakage: RR=0.3 (0.04, 3.2)	NR (RCT)
						Last 6 mo; 4-5 yr post randomization Erectile dysfunction: RR=1.8 (1.3, 2.6) Weak urinary stream: RR=0.7 (0.4, 1.2) Urinary incontinence: RR=1.7 (1.0, 2.8)	

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
						Last 6 mo; 6-8 yr post randomization Erectile dysfunction: RR=1.5 (1.2-2.0) Weak urinary stream: RR=0.5 (0.3-0.9) Urinary incontinence: RR=2.3 (1.3-3.9) Fecal leakage: NR [RR>1 indicates that symptoms were more common in the RP arm]	
Elliott[218] 2007 17570425 Retrospective cohort	CaPSURE	RP vs. WW	Treatment for urethral stricture*	Median 2.7 yr (range 3 days to 10.9 yr)	RP: 3310 WW: 378	HR 10.44 (3.28, 33.27), p<0.001	Age at treatment, clinical T stage, Gleason score, PSA at diagnosis, clinical risk stratification, BMI, urinary condition history, comorbidity count, race, marital status, education, household income

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Berge[231] 2007 17178188 Retrospective cohort	SEER-Medicare	WW vs. RP[a]	Cystoscopy; bladder irrigation/ cystostomy; TURP/bladder-neck incision; urethra dilation [procedures considered indicative of treatment-related morbidity]	5 yr	WW: 3612 RP: 3940	HR=1.00 (0.88, 1.13) for cystoscopy; P=0.071 for the null hypothesis that the coefficients of all treatments entered in the model are 0 HR=1.71 (1.33-2.20) for bladder irrigation/ cystostomy; P<0.001 for the null hypothesis that the coefficients of all treatments entered in the model are 0 HR=2.63 (2.08, 3.33) for TURP/bladder-neck incision; P =0.008 for the null hypothesis that the coefficients of all treatments entered in the model are 0 HR=0.71 (0.61, 0.84) for urethra dilation; P =0.309 for the null hypothesis that the coefficients of all treatments entered in the model are 0	Age, grade, comorbidity index
Quality of life							
Johansson[214] 2009 18783877 RCT	Ancillary investigation from SPCG-4 trial	WW vs. RP	77 questions on quality of life (developed based on interviews, tested for face validity on 30	Mean followup: 4.1 yr (range: 1-8)	RP: 189 (166 responders) WW: 187 (160 responders) [Some of the responders	Last 6 mo; 2-3 yr post randomization Anxiety (moderate or high): RR=0.5 (0.2, 1.0) Depressed mood (moderate or high): RR=0.6 (0.3, 1.0) Sense of well-being (low or	NR (RCT)

[a] In multivariable analysis this study used RP as the baseline treatment, thus adjusted estimates were reported for the comparison of RP with each other treatment (i.e., WW, RT, ADT). For the comparison of WW with other active treatments (i.e., WW vs. RT and ADT vs. RT) only unadjusted estimates were reported in the paper and were not extracted here. For more details please see the Methods section.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
			men, and validated in a "small pilot study"); trait measure from the Speilberger State-Trait Anxiety Inventory; Center for Epidemiological Studies Measure of Depression; psychological symptoms (anxiety, depressed mood) on 7-point VAS		did not reply to all questions]	moderate): RR=0.7 (0.4, 1.2) Self-assessed QoL (low to moderate): RR=1.0 (0.6, 1.6) Last 6 mo; 4-5 yr post randomization Anxiety (moderate or high): RR=1.2 (0.6, 2.4) Depressed mood (moderate or high): RR=1.5 (1.0, 2.6) Sense of well-being (low or moderate): RR=1.2 (0.7, 2.1) Self-assessed QoL (low to moderate): RR=1.0 (0.7, 1.6) Last 6 mo; 6-8 yr post randomization Anxiety (moderate or high): RR=0.6 (0.3, 1.1) Depressed mood (moderate or high): RR=0.8 (0.5, 1.3) Sense of well-being (low or moderate): RR=1.0 (0.6, 1.6) Self-assessed QoL (low to moderate): RR=0.8 (0.5, 1.1) [RR>1 indicates that a psychological symptom was more common in the RP arm]	

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Schapira[226] 2001 11242319 Prospective cohort	4 academically affiliated Wisconsin hospitals, including 2 VA Medical Centers	RP vs. EM	Disease-specific QoL: UCLA Prostate cancer Index General QoL: SF-36 scores	1 yr	RP: 37 EM: 25	Change in disease-specific QoL from pretreatment: Urinary function: RP = -27.8 vs. EM = +4.8 (P=0.004) Sexual function: RP = -38.4 vs. EM =-8.9 (P=0.01) Smaller (more negative) values indicate bigger reductions in QoL No significant difference between RP and EM groups in change in urinary bother, sexual bother, bowel function, or bowel bother index. Change in general QoL from pre-treatment – No significant difference between RP and EM groups in any domain.	Patient age, comorbidity, TNM, PSA, race, marital status, working status, and years of education.
Litwin[229] 2002 12115317 Retrospective cohort	CaPSURE	WW vs. RP	SF-36 scores at 24 months	1.5 yr	RP: 282 WW: 66	Mental domain: 85 ± 1.0 vs. 81 ± 2.4 Role of limitations due to emotional problems domain: 94 ± 2.0 vs. 86 ± 4.7 Vitality domain: 73 ± 1.4 vs. 66 ± 3.1 Social function domain: 100 ± 1.4 vs. 89 ± 2.2 (P < 0.05)	Comorbidity count, PSA at diagnosis, Gleason score on biopsy, age at the end of treatment

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Lubeck[227] 1999 9886609 Prospective cohort	CaPSURE	Observation vs. RP	HRQoL questionnaire composed of items from the RAND-36 item Health Survey. All scales were scored 0 to 100; optimal health or functioning = 100.	Immediate post-treatment period regression estimates	Observation: 87 RP: 351	Estimates are means (SE) based on linear mixed model for Observation vs. RP: Physical functioning, 74.9 (2.6); 73.5 (1.3) Role-physical, 61.7 (4.5); 34.3 (2.3) Role-emotional, 72.6 (3.9); 68.6 (2.0) Emotional well-being, 76.4 (1.7); 79.3 (0.9) Bodily pain, 73.4 (2.4); 73.0 (1.3) Social functioning, 80.9 (2.5); 67.2 (1.3) Energy/fatigue, 61.5 (2.2); 58.4 (1.1) General health perception, 60.4 (2.3); 76.0 (1.2) Health change, 38.1 (2.7); 42.5 (1.4) Health distress, 77.3 (2.1); 76.3 (1.1) Self-esteem, 82.1 (1.9); 82.4 (1.0)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RP	HRQoL questionnaire composed of items from the RAND-36 item Health Survey. All scales were scored 0 to 100; optimal health or functioning = 100.	1 yr post-treatment regression estimates	Observation: 87 RP: 351	Estimates are means (SE) based on linear mixed model for Observation vs. RP: Physical functioning, 72.7 (2.5); 91.9 (1.2) Role-physical, 60.2 (4.3); 94.0 (2.1) Role-emotional, 80.0 (3.8); 92.7 (1.8) Emotional well-being, 76.2 (1.6); 85.7 (0.8) Bodily pain, 75.9 (2.4); 91.7 (1.1) Social functioning, 80.4 (2.5); 97.8 (1.2) Energy/fatigue, 56.1 (2.1); 72.9 (1) General health perception, 60.2 (2.3); 76.7 (1.1) Health change, 54.3 (2.6); 69.6 (1.2) Health distress, 81.6 (2.0); 91.0 (1.0) Self-esteem, 78.2 (1.9); 84.2 (0.9)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RP	HRQoL questionnaire composed of items from the RAND-36 item Health Survey. All scales were scored 0 to 100; optimal health or functioning = 100.	2 yr post-treatment regression estimates	Observation: 87 RP: 351	Estimates are means (SE) based on linear mixed model for Observation vs. RP: Physical functioning, 70.8 (5.5); 85.8 (2.3) Role-physical, 62.8 (9.1); 72.3 (4.0) Role-emotional, 72.5 (8.0); 84.4 (3.4) Emotional well-being, 75.9 (3.4); 86.0 (1.5) Bodily pain, 76.1 (4.9); 84.4 (2.2) Social functioning, 77.9 (5.1); 88.8 (2.3) Energy/fatigue, 57.1 (4.4); 70.9 (2.0) General health perception, 53.8 (4.8); 74.6 (2.1) Health change, 46.4 (5.5); 60.2 (2.4) Health distress, 85.0 (4.2); 91.3 (1.9) Self-esteem, 88.2 (4.0); 87.7 (1.7)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RP	Prostate cancer-specific QoL items. All scales were scored 0 to 100; optimal health or functioning = 100.	Immediate post-treatment period regression estimates	Observation: 87 RP: 351	Estimates are means (SE) based on linear mixed model for Observation vs. RP: Urinary function, 81.9 (2.5); 43.4 (1.3) Bowel function, 82.5 (1.9); 78.1 (1.0) Sexual function, 37.0 (2.4); 10.3 (1.3) Urinary bother, 66.4 (3.3); 48.0 (1.7) Bowel bother, 81.8 (2.8); 64.7 (3.1) Sexual bother, 45.2 (4.3); 33.3 (2.3) CaP interference: family function, 81.1 (2.3); 81.9 (1.2) CaP interference, 83.0 (2.0); 65.8 (1.0)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RP	Prostate cancer-specific QoL items. All scales were scored 0 to 100; optimal health or functioning = 100.	1 yr post-treatment regression estimates	Observation: 87 RP: 351	Estimates are means (SE) based on linear mixed model for Observation vs. RP: Urinary function, 79.0 (2.4); 72.6 (1.2) Bowel function, 78.9 (1.9); 92.5 (0.9) Sexual function, 31.0 (2.4); 21.6 (1.1) Urinary bother, 70.4 (3.2); 82.1 (1.6) Bowel bother, 76.1 (2.7); 92.4 (1.3) Sexual bother, 42.6 (4.2); 36.7 (2.0) CaP interference: family function, 77.8 (2.2); 88.2 (1.1) CaP interference, 83.6 (2.0); 99.0 (0.9)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RP	Prostate cancer-specific QoL items. All scales were scored 0 to 100; optimal health or functioning = 100.	2 yr post-treatment regression estimates	Observation: 87 RP: 351	Estimates are means (SE) based on linear mixed model for Observation vs. RP: Urinary function, 87.4 (5.1); 70.7 (2.3) Bowel function, 89.1 (3.9); 88.1 (1.7) Sexual function, 29.1 (4.9); 26.8 (2.2) Urinary bother, 83.6 (6.7); 80.7 (3.0) Bowel bother, 90.0 (5.6); 90.2 (2.5) Sexual bother, 24.9 (8.8); 46.7 (3.9) CaP interference: family function, 86.9 (4.6); 88.6 (2.1) CaP interference, 87.0 (4.1); 90.6 (1.8)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.2. Comparison between watchful waiting and radical prostatectomy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Bacon[228] 2001 11586228 Prospective cohort	Health Professionals Followup Study	WW vs. RP	SF-36 summary measures and scales	Changes from 1996 to 1998	NR	Changes (SE) in SF-36 scores (RP as the baseline) Physical sum, -1.8 (2.8) Mental sum, 2.8 (3.3) Physical function, -3.7 (5.0) Role physical, -1.8 (12.1) Bodily pain, 1.3 (7.5) General health, -6.6 (5.7) Vitality, 5.6 (6.5) Social function, -0.1 (6.1) Role emotional, -1.4 (10.2) Mental health, 2.1 (4.3) None of the changes were statistically significant (P values: NR). The physical and mental health component scores were standardized to the US general population (mean score=50; SD=10)	Linear regression models adjusted for age category, SF-36 scale score before diagnosis and Gleason score.

CaP = carcinoma of the prostate; Dec = December; EM = expectant management; HR = hazard ratio; IPTW = inverse probability of treatment weights; IQR = interquartile range; IV = instrumental variable; mo = months; NR = not reported; NT = no treatment; POCS = Patterns of Care Study; PS = propensity score; PSA = prostate-specific antigen; QoL = quality of life; RP = radical prostatectomy; RR = relative risk; RT = radiation therapy; SD = standard deviation; SE = standard error; SMRW = standardized mortality ratio weights; TURP = transurethral resection of the prostate; VAS = visual analogue scale; WW = watchful waiting; yr = year

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Prostate cancer-specific mortality							
Stattin[223] 2010 20562373 Retrospective cohort	NPCRSFS	RT vs. surveillance	Death from prostate cancer as "underlying cause of death," data obtained from the Cause of Death Register or review of death certificates	Median followup= 8.2 yr (IQR=7.1-9.7 yr)	Surveillance: 2021 RT: 1429	HR=0.70 (0.45, 1.09)	Age at diagnosis, comorbidity, socioeconomic group, risk group.
All-cause mortality							
Wong[208] 2006 17164454 Retrospective cohort	SEER-Medicare	Radiation Tx vs. observation	Overall survival = interval from the date of diagnosis to the Medicare dare of death. Patients were censored at Dec. 20, 2002.	12 yr	RT: 18,249 Observation: 12,608	HR=0.81 (0.78, 0.85)	PS: age at diagnosis, SEER site, year of diagnosis, tumor size, tumor grade, marital status, residence in an urban setting, race, income, educational achievement, and 44 categorical variables encoding comorbidities. The authors reported a statistically significant interaction between tumor size and grade. For treatment subgroups (RP and radiation Tx) separate PS were built and used as covariates in the Cox regression models.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Stattin[223] 2010 20562373 Retrospective cohort	NPCRSFS	RT vs. surveillance	Death from any cause, data obtained from the Cause of Death Register or review of death certificates	Median followup= 8.2 yr (IQR=7.1-9.7 yr)	surveillance: 2021 RT: 1429	HR=0.68 (0.57, 0.82)	Age at diagnosis, comorbidity, socioeconomic group, risk group.
Elliott[218] 2007 17570425 Retrospective cohort	CaPSURE	BT vs. WW	Treatment for urethral stricture identified by study abstracted hospital records includes (ICD codes)*	Median 2.7 yr (range 3 days to 10.9 yr)	BT: 799 WW: 378	Crude stricture rates: 14/799 (1.8%) in patients received BT; 4/378 (1.1%) in patients received WW. HR=1.68 (0.46, 6.14), p=0.43	Age at treatment, clinical T stage, Gleason score, PSA at diagnosis, clinical risk stratification, BMI, urinary condition history, comorbidity count, race, marital status, education, household income
		EBRT vs. WW			EBRT: 645 WW: 378	Crude stricture rates: 11/645 (1.7%) in patients received EBRT; 4/378 (1.1%) in patients received WW. HR=1.77 (0.48, 6.55), p=0.39	Age at treatment, clinical T stage, Gleason score, PSA at diagnosis, clinical risk stratification, BMI, urinary condition history, comorbidity count, race, marital status, education, household income

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Quality of life							
Fransson[225] 2009 18985545 RCT	Ancillary investigation from UMEA1 trial	RT vs. WW	Prostate Cancer Symptom Scale; European Organization for Research and Treatment of Cancer QLQ-C30 (functioning and single symptom scales)	Median followup (range): WW, 114 mo (95-196); RT, 118 mo (77-188)	RT: 27 WW: 27	No statistically significant differences in health-related QoL function or symptom scales. No differences were detected between RT (mean=2.3) and WW (mean = 2.3; P=NR) for "limitations in daily life" caused by prostate cancer. No significant difference in "life situation" between RT (mean=3.7) and WW (mean=3.3; P=0.398). Mean scores on symptom and function scales at 10 years[b], WW vs. RT: Urinary bother: 2.0 vs. 2.8 (P=0.459) Incontinence: 1.0 vs. 1.5 (P=0.092) Urinary frequency/day: 6.1 vs. 6.5 (P=0.639) Nocturia: 2.1 vs. 1.8 (P=0.900) Starting problems: 1.7 vs. 1.9 (P=0.354) Pain while urinating: 0.2 vs. 0.4 (P=0.099) Urgency: 2.3 vs. 2.1 (P=0.750) Daily activity: 1.6 vs. 1.4 (P=0.984) Bowel bother: 1.6 vs. 2.7 (P=0.094) Daily activity: 0.9 vs. 1.3 (P=0.364)	For the mean scores univariate non-parametric (Mann-Whitney) tests were used (RCT). MANOVA analyses did not detect any associations between overall QoL/health, urinary bother, bowel bother and sexual bother (dependent variables) and age, RT dose (RT group only), progress (RT and WW), hormones (RT and WW), diabetes, or smoking habits.

[b] The study also reported mean symptom/function scores at 4 years of followup; we have extracted information only for the 10 year followup period.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
						Stool frequency: 1.5 vs. 2.0 (P=0.296) Blood in stools: 0.4 vs. 1.4 (P=0.085) Stool leakage: 1.0 vs. 1.2 (P=0.430) Bowel movements: 2.4 vs. 2.1 (P=0.660) Cramp: 0.0 vs. 0.1 (P=0.295) Planning: 1.0 vs. 1.5 (P=0.765) Mucus: 0.5 vs. 0.9 (P=0.103) Sexual bother: 3.8 vs. 7.4 (P=0.011) Desire: 6.4 vs. 8.0 (P=0.169) Erection: 7.2 vs. 8.4 (P=0.292) Weak urinary stream: 3.0 vs. 4.8 (P=0.034) No difference was observed at 10 years between RT and WW regarding the question “How would you feel if you lived the rest of your life with your urinary problems as they are now?” (P=0.643). No difference was observed at 10 years between RT and WW regarding the question “How would you feel if you lived the rest of your life with your bowel problems as they are now?” (P=0.653). No difference was observed at 10 years between RT and WW regarding the question “How	

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
						would you feel if you lived the rest of your life with your sexual problems as they are now?" (P=0.819).	
Litwin[229] 2002 12115317 Retrospective cohort	CaPSURE	WW vs. RT	Mean SF-36 scores at 24 months	1.5 yr	RT: 104 WW: 66	Mental domain: 75 ± 1.9 vs. 81 ± 2.4 Role of limitations due to emotional problems domain: 81 ± 3.8 vs. 86 ± 4.7 Vitality domain: 61 ± 2.5 vs. 66 ± 3.1 Social function domain: 86 ± 2.7 vs. 89 ± 2.2 (P < 0.05)	Comorbidity count, PSA at diagnosis, Gleason score on biopsy, age at the end of treatment
Schapira[226] 2001 11242319 Prospective cohort	4 academically affiliated Wisconsin hospitals, including 2 VA Medical Centers	RT vs. EM	Disease-specific QoL: UCLA Prostate cancer Index General QoL: SF-36 scores	1 yr	RT: 40 EM: 25	No significant difference between RT and EM groups in any domain in both disease-specific and general QoL measures.	Patient age, comorbidity, TNM, PSA, race, marital status, working status, and years of education.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Thong[224] 2009 19747357 Retrospective matched cohort	Eindhoven Cancer Registry (ECR)	RT vs. "AS"	General QoL: SF-36 scores Expanded Prostate Cancer Index (EPIC): urinary and bowel functioning, and urinary and bowel bother Quality of Life – Cancer Survivors (QOL-CS)	Mean 8 yr	RT: 71 AS: 71	RT was negatively associated with physical functioning and bodily pain dimensions of the SF-36, spiritual and total wellbeing scores of the QoL-CS, and bowel function and bowel bother of EPIC index. No other significant associations between general QoL, cancer-specific QoL, or disease-specific QoL scores and management strategy (RT vs. "AS").	Matching: cancer stage, tumor grade, age at diagnosis (±2 yrs), and number of years since diagnosis (±2 yrs). Multivariate model adjusted for comorbidity and disease progression

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Lubeck[227] 1999 9886609 Prospective cohort	CaPSURE	Observation vs. RT	HRQoL questionnaire composed of items from the RAND-36 item Health Survey. All scales were scored 0 to 100; optimal health or functioning = 100.	Immediate post-treatment period regression estimates	Observation: 87 RT: 75	Estimates are means (SE) based on linear mixed model for Observation vs. RT: Physical functioning, 74.9 (2.6); 76.5 (3.0) Role-physical, 61.7 (4.5); 51.0 (5.1) Role-emotional, 72.6 (3.9); 65.8 (4.5) Emotional well-being, 76.4 (1.7); 77.5 (1.9) Bodily pain, 73.4 (2.4); 76.4 (2.7) Social functioning, 80.9 (2.5); 76.0 (2.8) Energy/fatigue, 61.5 (2.2); 58.1 (2.5) General health perception, 60.4 (2.3); 65.3 (2.7) Health change, 38.1 (2.7); 38.1 (3.0) Health distress, 77.3 (2.1); 80.4 (2.3) Self-esteem, 82.1 (1.9); 83.3 (2.2)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RT	HRQoL questionnaire composed of items from the RAND-36 item Health Survey. All scales were scored 0 to 100; optimal health or functioning = 100.	1 yr post-treatment regression estimates	Observation: 87 RT: 75	Estimates are means (SE) based on linear mixed model for Observation vs. RT: Physical functioning, 72.7 (2.5); 76.7 (2.5) Role-physical, 60.2 (4.3); 66.3 (4.2) Role-emotional, 80.0 (3.8); 81.9 (3.7) Emotional well-being, 76.2 (1.6); 74.6 (1.6) Bodily pain, 75.9 (2.4); 75.6 (2.3) Social functioning, 80.4 (2.5); 84.6 (2.4) Energy/fatigue, 56.1 (2.1); 56.7 (2.1) General health perception, 60.2 (2.3); 62.1 (2.3) Health change, 54.3 (2.6); 61.0 (2.5) Health distress, 81.6 (2.0); 81.8 (2.0) Self-esteem, 78.2 (1.9); 79.8 (1.8)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RT	HRQoL questionnaire composed of items from the RAND-36 item Health Survey. All scales were scored 0 to 100; optimal health or functioning = 100.	2 yr post-treatment regression estimates	Observation: 87 RT: 75	Estimates are means (SE) based on linear mixed model for Observation vs. RT: Physical functioning, 70.8 (5.5); 65.1 (4.8) Role-physical, 62.8 (9.1); 55.4 (8.2) Role-emotional, 72.5 (8.0); 75.8 (7.3) Emotional well-being, 75.9 (3.4); 77.6 (3.1) Bodily pain, 76.1 (4.9); 73.8 73.8 (4.5) Social functioning, 77.9 (5.1); 76.9 (4.7) Energy/fatigue, 57.1 (4.4); 54.1 (4.0) General health perception, 53.8 (4.8); 53.9 (4.4) Health change, 46.4 (5.5); 47.3 (5.0) Health distress, 85.0 (4.2); 84.2 (3.9) Self-esteem, 88.2 (4.0); 82.2 (3.6)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RT	Prostate cancer-specific QoL items. All scales were scored 0 to 100; optimal health or functioning = 100.	Immediate post-treatment period regression estimates	Observation: 87 RT: 75	Estimates are means (SE) based on linear mixed model for Observation vs. RT: Urinary function, 81.9 (2.5); 79.8 (2.8) Bowel function, 82.5 (1.9); 69.2 (2.2) Sexual function, 37.0 (2.4); 34.4 (2.7) Urinary bother, 66.4 (3.3); 57.3 (3.7) Bowel bother, 81.8 (2.8); 77.9 (1.4) Sexual bother, 45.2 (4.3); 46.1 (5.0) CaP interference: family function, 81.1 (2.3); 82.4 (2.6) CaP interference, 83.0 (2.0); 78.4 (2.3)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.
		Observation vs. RT	Prostate cancer-specific QoL items. All scales were scored 0 to 100; optimal health or functioning = 100.	1 yr post-treatment regression estimates	Observation: 87 RT: 75	Estimates are means (SE) based on linear mixed model for Observation vs. RT: Urinary function, 79.0 (2.4); 80.7 (2.4) Bowel function, 78.9 (1.9); 83.0 (1.8) Sexual function, 31.0 (2.4); 34.1 (2.3) Urinary bother, 70.4 (3.2); 75.5 (3.2) Bowel bother, 76.1 (2.7); 76.6 (2.7) Sexual bother, 42.6 (4.2); 42.1 (4.2) CaP interference: family function, 77.8 (2.2); 84.3 (2.2) CaP interference, 83.6 (2.0); 86.2 (1.9)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

Appendix Table C4.3. Comparison between watchful waiting and radiation therapy (continued)

Author, Year [Pubmed ID] Study design	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
		Observation vs. RT	Prostate cancer-specific QoL items. All scales were scored 0 to 100; optimal health or functioning = 100.	2 yr post-treatment regression estimates	Observation: 87 RT: 75	Estimates are means (SE) based on linear mixed model for Observation vs. RT: Urinary function, 87.4 (5.1); 84.9 (4.7) Bowel function, 89.1 (3.9); 83.1 (3.5) Sexual function, 29.1 (4.9); 25.4 (4.5) Urinary bother, 83.6 (6.7); 65.1 (6.2) Bowel bother, 90.0 (5.6); 74.8 (5.3) Sexual bother, 24.9 (8.8); 32.4 (8.1) CaP interference: family function, 86.9 (4.6); 84.7 (4.2) CaP interference, 87.0 (4.1); 83.1 (3.7)	Treatment group interaction with time, age, time of treatment. P values for comparisons between treatment groups NR.

ADT = androgen deprivation therapy; BT = brachytherapy; HR = hazard ratio; MANOVA = multivariate analysis of variance; NR = not reported; RP = radical prostatectomy; RT = radiation therapy.

Appendix Table C4.4. Comparison between watchful waiting and other active treatments

Author Year Pubmed id	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Prostate cancer-specific mortality							
Wong[208] 2006 17164454 Retrospective cohort	SEER-Medicare	Active treatment (RP or BT or RT considered in aggregate) vs. observation	Death from prostate cancer based on the cause of death reported in SEER. Data on cause-specific mortality were available through the end of 2000.	12 yr	Active treatment: 32,022 Observation: 12,608	HR=0.67 (0.58, 0.77)	PS: age at diagnosis, SEER site, year of diagnosis, tumor size, tumor grade, marital status, residence in an urban setting, race, income, educational achievement, and 44 categorical variables encoding comorbidities. The authors reported a statistically significant interaction between tumor size and grade. For the association of treatment and survival in the entire cohort, estimates were adjusted for the PS and comorbidities, tumor grade, and tumor size as categorical variables.
All-cause mortality							
Wong[208] 2006 17164454 Retrospective cohort	SEER-Medicare	Active treatment (RP or BT or RT considered in aggregate) vs. observation	Overall survival = interval from the date of diagnosis to the Medicare dare of death. Patients were censored at Dec. 20, 2002.	12 yr	Active treatment: 32,022 Observation: 12,608	HR=0.69 (0.66, 0.72); stratified by PS quintile, HR 0.67 (0.65-0.70)	PS: age at diagnosis, SEER site, year of diagnosis, tumor size, tumor grade, marital status, residence in an urban setting, race, income, educational achievement, and 44 categorical variables encoding comorbidities. The authors reported a statistically significant interaction between tumor size and grade. For the association of treatment and survival in the entire cohort, estimates were adjusted for the PS and comorbidities, tumor grade, and tumor size as categorical variables.

Appendix Table C4.4. Comparison between watchful waiting and other active treatments (continued)

Author Year Pubmed id	Study name /Database	Comparison	Outcome definition/ measurement instrument	Followup (yr)	Sample size per group	Results	Factors included in the model
Morbidity of primary treatment							
Elliott[218] 2007 17570425 Retrospective cohort	CaPSURE	RP+EBRT (combination treatment) vs. WW	Treatment for urethral stricture identified by study abstracted hospital records includes (ICD codes)*	Median 2.7 yr (range 3 days to 10.9 yr)	RP+EBRT: 73 WW: 378	Crude stricture rates: 2/73 (2.7%) in patients received RP+EBRT; 4/378 (1.1%) in patients received WW. HR=4.39 (0.72-26.69), p=0.11	Age at treatment, clinical T stage, Gleason score, PSA at diagnosis, clinical risk stratification, BMI, urinary condition history, comorbidity count, race, marital status, education, household income
		BT+EBRT (combination treatment) vs. WW			BT+EBRT: 231 WW: 378	Crude stricture rates: 12/231 (5.2%) in patients received BT+EBRT; 4/378 (1.1%) in patients received WW. HR=4.56 (1.23-16.88), p=0.02	Age at treatment, clinical T stage, Gleason score, PSA at diagnosis, clinical risk stratification, BMI, urinary condition history, comorbidity count, race, marital status, education, household income

Appendix Table C4.5: Cost comparisons of watchful waiting with active treatments

Author year UI	Design	Data source	Years	N Population characteristics	Methods	Costs	WW (95% CI)	RP (95% CI)	RT (95% CI)	Notes
Snyder[159] 2010 20734396	Retrospective cohort study (also includes a comparison with a non-cancer control group)	SEER-Medicare	2000, followed for 5 yr	WW = 2805 RP monotherapy = 2200 RT monotherapy = 2582 Total N = 13,769 (included 3992 patients receiving hormonal therapy + RT or medical pADT) ≥65 yr; lived in SEER regions, clinically localized prostate cancer Control group (n=13,769; matched for age, sex, race, region, comorbidity, survival)	Costs from Medicare adjusted for inflation to 2007 dollars and geographic region; discounted 3% per yr after 1st yr; grouping based on treatment received 1st 9 mo; estimates were derived from IPTW analysis using a propensity score accounting for age, race, comorbidity, SEER region, urban/rural location, socioeconomic status and grade	Incremental costs vs. control: Yr 1 Total over 5 yr	 $ 3936 (3078-4794) $ 8535 ($ 6223-$ 10,847)	RP monotherapy $ 15,556 (14,835-16,277) $ 19,481 ($ 17,538-$ 21,424)	RT monotherapy $ 12,319 (11,419-13,219) $ 16,653 ($ 14,228-$ 19,078)	Required patients to have survived at least 9 mo post diagnosis; total costs were only calculated for years during which the patient survived.

Appendix Table C4.5: Cost comparisons of watchful waiting with active treatments (continued)

Author year UI	Design	Data source	Years	N Population characteristics	Methods	Costs	WW (95% CI)	RP (95% CI)	RT (95% CI)	Notes
Andersson[237] 2011 21265595	Substudy of RCT	SPCG-4	Recruitment, 1989-99; followed through July 2007	WW = 105 RP = 107 <75 years, life expectancy >10 years, T0d-T2 disease, WHO well/moderately differentiated, PSA <50ng/ml, no evidence of skeletal metastases on bone scan; patients from the trial were included if they resided in the counties where the two centers that randomized most patients were located (Örebro and Uppsala).	Medical records were retrospectively reviewed; a healthcare provider perspective was adopted (costs generated from care outside the hospital were not considered); resource use was measured in physical units and then multiplied by the unit cost (based on 2007 Swedish prices converted to €).	Total mean cost after median followup of 11.8 yr for the WW group and 12.2. for the RP group	€ 18,124 (NR)	€ 24,247 (NR)	NA	No patient was lost to follow-up P<0.01 for the absolute difference between groups, unadjusted analysis; in multivariable analysis P=0.003 (adjusted for age, PSA and Gleason score).
Penson[235] 2001 11248628	Retrospective cohort study	CaPSURE	1990-97	WW = 37 Active Tx = 198 [RP monotherapy, RP + neoadjuvant hormone therapy; RT monotherapy, RT + neoadjuvant hormonal therapy;	Direct costs of prostate cancer treatment (outpatient visits, laboratory tests and procedures, prescribed medications, surgical treatments and hospitalizations), adjusted to 1996	Average first year cost	$ 484 (NR)	RP monotherapy $ 7320 (NR)	RT monotherapy $ 7430 (NR)	P <0.001 for the treatment effect (ANCOVA, adjusted for stage, Gleason sum, serum PSA at diagnosis, insurance status, comorbidities, age); cost

Appendix Table C4.5: Cost comparisons of watchful waiting with active treatments (continued)

Author year UI	Design	Data source	Years	N Population characteristics	Methods	Costs	WW (95% CI)	RP (95% CI)	RT (95% CI)	Notes
				medical pADT, orchiectomy, medical pADT followed by orchiectomy]	values; outpatient costs were calculated using 1996 Medicare Physician and Laboratory Fee Schedule; outpatient facility costs per unit were obtained from average Medicare service costs (1992); total costs were calculated by multiplying the frequency of each service by the unit cost; hospitalization costs were obtained from Medicare (1994); drug costs were obtained from the 1996 Red Book					estimates are unadjusted means.

Appendix Table C4.5: Cost comparisons of watchful waiting with active treatments (continued)

Author year UI	Design	Data source	Years	N Population characteristics	Methods	Costs	WW (95% CI)	RP (95% CI)	RT (95% CI)	Notes
Wilson[a236] 2007 17186528	Retrospective cohort study	CaPSURE	1995-2004	WW = 238 RP = 2496 BT = 668 EBRT = 409 ADT = 607 Cryotherapy = 135 [We did not extract data on the costs of cryotherapy or ADT]	Recurrent event analysis to calculate mean cumulative functions of prostate-related costs	Mean cumulative function cost over 5.5 yr Total Medication costs Office visit costs Hospitalization costs *[Low risk patients]*	$ 31,871 $ 5269 $ 19,716 $ 6886	$ 32,795 $ 2832 $ 18,116 $ 11,847	$ 28,366 [BT] $ 48,840 [EBRT] $ 3741 [BT] $ 6515 [EBRT] $ 18,844 [BT] $ 36,446 [EBRT] $ 5780 [BT] $ 5879 [EBRT]	No statistical comparison was reported for these analyses. (the authors claim to have "stratified" or "controlled" for patient age, risk group, and ethnicity; it is not clear if the extracted estimates were adjusted for these variables)

[a] The study included an unknown number of T3a patients. Because such patients were considered "high risk", we have only extracted information on the subgroups of patients classified as "low" or "intermediate" risk.

Appendix Table C4.5: Cost comparisons of watchful waiting with active treatments (continued)

Author year UI	Design	Data source	Years	N Population characteristics	Methods	Costs	WW (95% CI)	RP (95% CI)	RT (95% CI)	Notes
						Mean cumulative function cost over 5.5 yr				Same as above
						Total	$ 31,789	$ 35,037	$ 41,419 [BT] $ 56,725 [EBRT]	
						Medication costs	$ 5098	$ 6869	$ 4940 [BT] $ 9164 [EBRT]	
						Office visit costs	$ 21,757	$ 18,180	$ 23,873 [BT] $ 43,327 [EBRT]	
						Hospitalization costs	$ 4934	$ 9989		
						[Intermediate risk patients]			$ 12,605 [BT] $ 4234 [EBRT]	

Appendix Table C5.1. Proportion of patients on AS who went on to receive active treatments[a]

AS Cohort	Author, yr PMID	Median follow up (range)	No. on AS	No. received treatment	% who received active treatment
Toronto	Klotz, 2010 19917860 Krakowsky, 2010 20478589	6.8 yr (1-13 yr)	450	135	30 (10 yr actuarial 38%)
Johns Hopkins	Ross, 2010 20439642	2.9 yr (0.48-12.2 yr)	290	102	35
U. of Miami	Soloway, 2010 20800964	3.7 yr (mean)	230	32	14
Kagawa U.	Kakehi, 2008 18272471	up to 4.5 yr	118	64	54 (3 yr actuarial 51.1%)
McGill U.	Al Otaibi, 2008 18484590	6.3 yr (1.7-14 yr)	92 had ⩾1 rebiopsy	29[b]	32
PRIAS	van den Bergh, 2010 19817747	1.02 yr (0.6-1.5 yr)	500	82	16
Royal Marsden	van As, 2009 19095345	2.4 yr (mean)	86 had ⩾1 rebiopsy	39	45
UCSF	Cooperberg, 2011 21115873	up to 4 yr	376	119[c]	32
4 centers (Miami, UBC, Sloan-Kettering, Cleveland Clinic)	Eggener, 2009 19233410	2.4 yr (IQR 1.2, 4.3 yr)	262	43	16
Dana Farber	San Francisco, 2011 21167525	2.4 yr	120 had ⩾1 rebiopsy	NR	NR
U. of Connecticut	Ercole, 2008 18707696	4 yr (1-14 yr)	40	9[d]	23
Baylor or Sloan-Kettering	Patel, 2004 15017211	3.7 yr	88	31[e]	35
Sloan-Kettering	Adamy, 2011 21167529	to 5 yr	238	NR	NR
Cleveland Clinic	Miocinovic, 2011 21256549	2.8 yr (IQR 1.7, 3.8 yr)	116 had ⩾1 rebiopsy	38	33
Canary PASS (Stanford, UCSF, UBC, U of Washington, U of Texas at San Antonio)	Newcomb, 2010 19758683	recruiting	NR	NR	NR
ProtecT (multicenters in UK)	Metcalfe, 2009 19603015	Followup phase	408 (patients who select AS actively: observational cohort) 138 (randomized to AS in the RCT)	NR	NR

AS = active surveillance; IQR = interquartile range; NR = not reported; U. = university; yr = year.

[a] Data from papers eligible for inclusion in this report.
[b] 6 had ADT alone
[c] 14 had ADT
[d] 2 had ADT
[e] 1 had ADT

U.S. Department of
Health and Human Services

Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850

AHRQ Pub. No. 12-E003-EF
December 2011