Draft Evidence Report
Number XXX

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.

Prepared by:

Redacted

Investigators:

AHRQ Publication No. 0X-E0XX
Publication Month and Year

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted for which further reproduction is prohibited without the specific permission of copyright holders.

Suggested Citation:

No investigators have any affiliations or financial involvement (e.g., employment, consultancies, honoraria, stock options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in this report.

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. The Office of Dietary Supplements/National Institutes of Health, the Public Health Agency of Canada, Health Canada, and Food and Drug Administration requested and provided funding for this report. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ 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.

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 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

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

Stephanie Chang, M.D., M.P.H.
Director, EPC Program
Agency for Healthcare Research and Quality

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

Acknowledgments

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: Systematically review the role of active surveillance for triggers to begin curative treatment in men with low-risk prostate cancer. Key Questions address the change 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, and existing systematic and narrative reviews.

Study selection: Randomized controlled trials and nonrandomized comparative studies of treatments, and multivariable association studies. 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, outcomes, and study methodological issues.

Data synthesis: In total, 169 articles met eligibility criteria (65 on trends, 49 on definitions, 37 on factors, 18 on comparative effectiveness). Increased diagnosis of early-stage prostate cancer led to an observed 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. Over time, a smaller proportion of men were on observational managements versus active treatments, even among those with low-risk disease. There is no standardized definition of active surveillance. Fifteen cohorts used different monitoring protocols, using different combinations of periodic digital rectal examination, prostate-specific antigen (PSA) testing, rebiopsy, and/or imaging findings. Predictors that a patient receives no initial active treatment included older age, presence of comorbidities, higher Gleason score, higher tumor stage, higher diagnostic PSA, higher disease progression risk group, and decreased baseline anxiety. No trial provided results comparing men with localized disease on active surveillance with surgery or radiation therapy.

Limitations: Because of the different usages of the terms “active surveillance” and “watchful waiting” and their intended and often mixed (both curative and palliative) treatment objectives, it is difficult to determine which patients in the studies had active monitoring for triggers indicative of curative treatment or 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 in this field. Ongoing clinical trials may provide information on the comparative effectiveness of active surveillance compared to immediate active treatment, but will require long term followup.

Table of Content

Preface ..... **iv**

Acknowledgments ..... **v**

Structured Abstract ..... **vi**

Table of Content ..... **vii**

Abbreviations and acronyms ..... **x**

Executive Summary **1**
- Background 1
- Objectives 2
- Methods 4
- Results 6
- Discussion 17

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
- Data Extraction and Summaries 11
- Quality Assessment 12
- Data synthesis and presentation 13
- Summary Tables 13
- Grading the Body of Evidence for Key Question 4 14
- Peer Review and Public Commentary 15

Results **17**
- 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? 18
- Key question 2: How are active surveillance and other observational management strategies defined? 30
- Key Question 3. What factors affect the offer of, acceptance of, and adherence to active surveillance? 61
- 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? 73
- Key Question 5. What are the research needs regarding active surveillance (or watchful waiting) in localized prostate cancer? 95

Discussion **101**

Appendix A. Search strategies **A-1**

Appendix B. Ongoing randomized studies comparing observational management strategies with active treatment strategies for the treatment of clinically localized disease B-1

Appendix C. Appendix Tables of Key Question 1 to Key Question 4 ..... C-1

Tables and Figures

Figure 1. Analytic framework that depicts the five Key Questions (KQ) examining the role of active surveillance in the management of men with clinically localized prostate cancer 6

Figure 2. Literature flow 17

Figure 1.1. Years covered and databases utilized by studies considered eligible for Key Question 1 20

Figure 1.2. Age-adjusted SEER incidence rates for prostate cancer (1975-2008) 21

Figure 2.1. Summary of 12 cohorts that used Gleason score as part of AS program eligibility criteria 32

Figure 2.2. Summary of 10 cohorts that used PSA (ng/mL) as part of AS program eligibility criteria 33

Table 2.2. Eligibility criteria for enrollment in protocols with curative intent in chronological order of starting enrollment year 34

Table 2.3. Monitoring criteria in protocols with curative intent in chronological order of starting enrollment year 37

Table 2.4. Followup frequencies of 15 unique AS cohorts[a] 42

Figure 2.3. Summary of 7 cohorts that used PSA (ng/mL) or PSA kinetics as part of trigger for curative treatment 43

Table 2.5. Unique 13 cohorts of observational management strategies with palliative intent 44

Table 2.6. Eligibility criteria for enrollment in protocols with palliative intent in chronological order of starting enrollment year 48

Table 2.7. Monitoring criteria in protocols of observational management strategies with palliative intent in chronological order of starting enrollment year 50

Table 2.10. Protocols that did not report information on triggers for intervention in chronological order of starting enrollment year 57

Table 2.11. Monitoring parameters in cohorts that did not report information on triggers for intervention in chronological order of starting enrollment year 58

Table 4.1. Descriptive characteristics of the randomized controlled trials and comparative cohort studies considered relevant to KQ4 77

Table 4.2. Comparison between active surveillance or watchful waiting and radical prostatectomy 83

Table 4.3. Comparison between active surveillance or watchful waiting and radiation therapy 88
Table 4.4. Comparison between active surveillance or watchful waiting and other active treatments 91

Abbreviations and acronyms

2D-RT	conventional (non-3D planned) RT
3D-CRT	conformal RT
ADT	androgen deprivation therapy
AHRQ	Agency for Healthcare Research and Quality
AS	active surveillance
AUA	American Urological Association
BT	Brachytherapy
CaPSURE	Cancer of the Prostate Strategic Urologic Research Endeavor
CDP	Consensus Development Program
EBRT	external beam RT (including 2D-RT, 3D-CRT, IMRT)
EPC	Evidence-based Practice Center
EPIC	Expanded Prostate Cancer Index Composite
Gy	Gray (unit of radiation dose)
ICER	Institute for Clinical and Economic Review
IGRT	image guided RT (including IMRT and SBRT)
IMRT	intensity modulated RT
NCCN	National Comprehensive Cancer Network
NCDB	National Cancer Database
NCHS	National Center for Health Statistics
NCI	National Cancer Institute
NIH	National Institutes of Health
PSA	Prostate-specific antigen
QoL	quality of life
RT	radiation therapy (radiotherapy)
RTOG	Radiation Therapy Oncology Group
SEER	Surveillance, Epidemiology, and End Results Program of the NCI
TURP	transurethral resection of the prostate
TRUS	Transrectal ultrasound
WW	watchful waiting
PCOS	Prostate Cancer Outcomes Study
PIVOT	Prostate cancer Intervention Versus Observation Trial
POCS	Patterns of Care Study
ProtecT	Prostate Testing for Cancer and Treatment trial
QoL	Quality of life
QoL-CS	Quality of Life-Cancer Survivors
TEP	Technical Expert Panel
TOO	Task Order Officer

Executive Summary

Background[a]

In 2011, over 240,000 men are projected to be diagnosed with prostate cancer and 33,000 to die from the condition. Currently, in the United States, most instances 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. Notably, though, 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 much likelihood of obtaining any clinical benefit due to the slow progression of many prostate tumors. However, 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 generally connotes the monitoring of a potentially curable prostate cancer and intervening with a curative-intent treatment at the earliest sign of worrisome progression. In contrast, WW generally connotes postponing therapeutic interventions until symptom development, with the primary objective being palliation of the symptoms rather than an attempt at a cure. 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 a relatively passive strategy—with interventions triggered by symptoms. It should be underscored, however, that in the scientific literature the two terms and their intents are often used interchangeably.

Given the tradeoffs between complications from curative treatments and long-term risks of delaying treatment, and thus the use of AS and other observational management strategies by men who are more interested in avoiding the risks of curative treatment, it is important to clarify appropriate eligibility criteria and followup protocols for the 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. It is also of interest to evaluate whether men offered AS will accept this strategy and adhere to it. If men feel a strong need “to do something” to definitively treat the cancer, and thus AS is rarely chosen or not adhered to, then the impact of offering this strategy will be small. Therefore, factors that relate to the offer of AS by clinicians to patients, acceptance of AS by patients and their families, and adherence with AS once this course has been chosen need exploration.

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 the present evidence review for use in this conference.

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

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 (CDP), the purpose of which is to evaluate the scientific evidence on a particular topic and develop a consensus 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 the present evidence review, and will then use that information to compose their statement. Additional information about the NIH Consensus Development Program (CDP) 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 to systematically review the literature to address the conference questions. Key Question 1 pertains to temporal trends in the natural history of prostate cancer in the U.S. 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, such observational management strategies. Key Question 4 pertains to the comparative effectiveness of treatments for localized prostate cancer. And 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 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. Follow up 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 – 2nd 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?

Analytic Framework

To guide this systematic review and facilitate the interpretation of Key Questions, we developed an analytic framework (Figure 1) that depicts the logical progression and interconnection of all five Key Questions for this report.

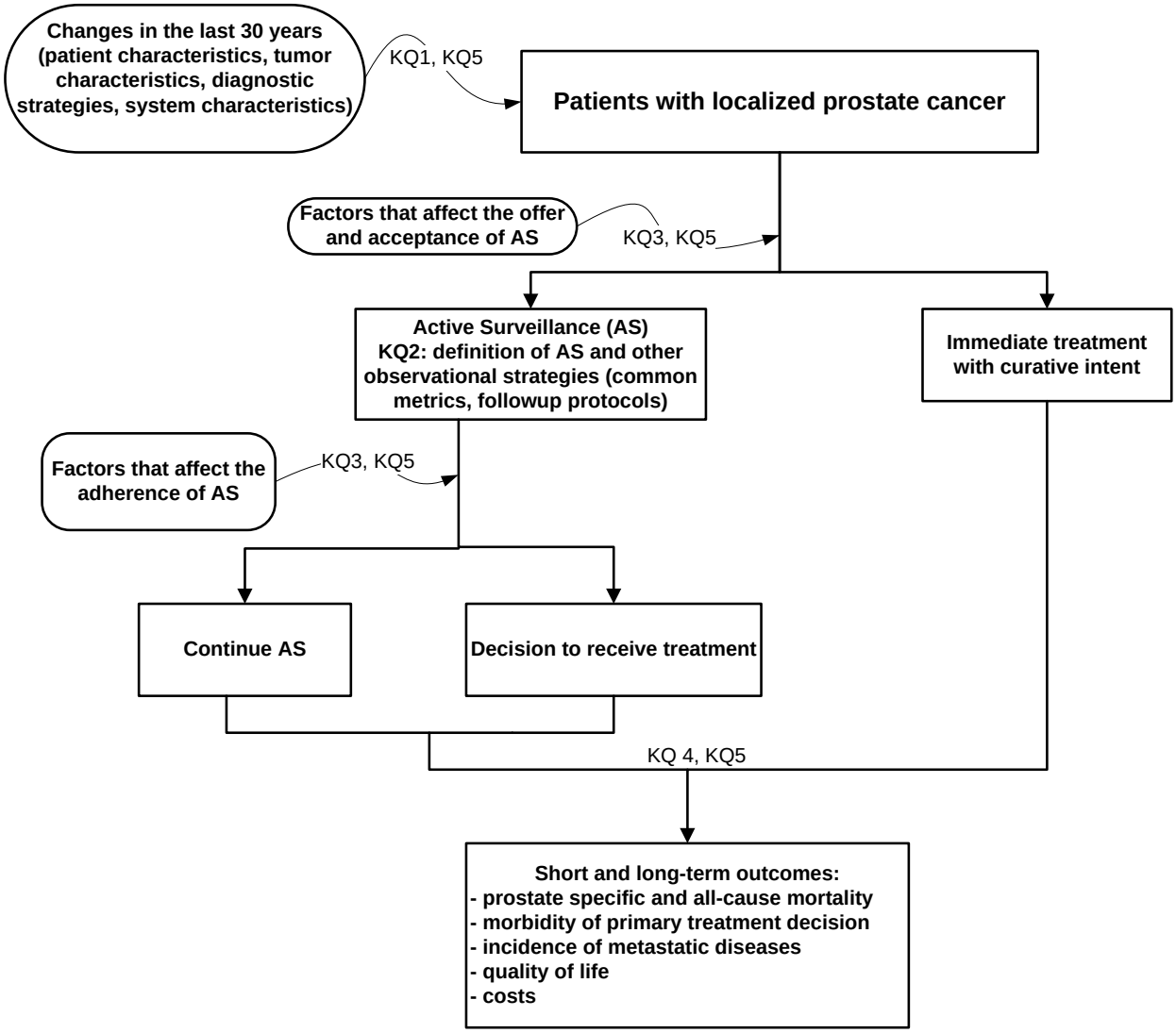


Figure 1. Analytic framework that depicts the five Key Questions (KQ) that examine the role of active surveillance in the management of men with clinically localized prostate cancer

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 March 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 topics. 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 primarily relied on previous systematic reviews on prostate cancer conducted for the AHRQ EPC program. These

searches were supplemented with studies recommended to us by the Technical Expert Panel and from reference lists of eligible primary studies and relevant review articles. We did not include unpublished data.

The following are the study eligibility criteria 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. Minimum sample size of 1000 patients.

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. First, we included studies that used quantitative methods to analyze databases or cohorts of patients to elucidate predictors of the offer of, 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. Second, we included 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 must have used a predefined approach to collect information. Third, we also searched for experimental studies evaluating the effect of tools such as decision aids on offer, acceptance, or adherence of 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 diseases. 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, satisfaction with treatment, and costs.

All five EPC team members participated in screening and selecting studies. An iterative process was used to ensure training and 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

For all studies, we extracted bibliographic data, eligibility criteria, enrollment years, study duration, and sample size. 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 that had 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 the methodological quality only for studies included for Key Question 4. Studies were graded using standard AHRQ EPC methodology with a three-category 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 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. Quality A studies have the least likelihood of bias and are considered most 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 appendices) 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 to 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 64 relevant primary observational studies and one systematic review. Of the primary observational studies, 42 analyzed the Surveillance Epidemiology and End Results (SEER) database of the National Cancer Institute (NCI) or a subset of its component registries, six additional studies utilized the linked SEER-Medicare database, ten the Cancer of the Prostate Urologic Research Endeavor (CaPSURE) database, four the National Cancer Database (NCDB), and two analyzed other large U.S.-based databases. In addition we queried the online SEER database.

Trends in prostate cancer incidence

Prostate cancer incidence rose between 1975 and 1992 and then fell until around 1995. After a period of nonsignificant increase from 1995 to 2000, rates declined again from 2000 to 2007. Overall, 28 studies provided information on trends of prostate cancer incidence.

Age

Ten studies (covering 1973-2005) 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, 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

In 14 studies (covering 1973-2005) all racial/ethnic 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

Thirteen studies (covering 1969-2005) 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. These studies also consistently demonstrated a decrease in incidence rates for all disease stages from mid-1990s to 2000. A single study investigated changes in the distribution of T stage over time 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 and 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.

Tumor grade

Six studies (covering 1973-2005) stratified prostate cancer incidence by tumor grade (level of differentiation or Gleason score). In these studies, 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).

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. Overall, 17 studies provided information on trends of prostate cancer mortality or survival after diagnosis.

Age

Eight studies (covering 1969 to 2003) demonstrated decreases in the mortality rate for all age groups between the early-1990s and 1999. One study of prostate cancer-specific survival demonstrated that over time (1988-95) 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).

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

Thirteen studies (covering 1969-2000) demonstrated that mortality rates among blacks were consistently higher compared to that of non-Hispanic whites in all studies and across time periods. Similarly, 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.

Tumor stage

Three studies (covering 1969-99) 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.

Tumor grade

One of two studies (of patients diagnosed in 1973-95) 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 differentiated cancers and poorly differentiated disease 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).

Patient, tumor and system-level characteristics at diagnosis

We identified 43 observational studies reporting on patient characteristics at presentation.

Age

Eighteen studies (covering 1973-2005) reported on patients' age at presentation. Among six studies evaluating average age at diagnosis of prostate cancer, four found reductions in the average age of patients whereas two studies did not report any changes during their respective time periods. The 12 studies that evaluated distribution of patients' ages generally supported a trend toward younger age at diagnosis (the effect was significant in four of the six studies reporting statistical tests).

Comorbidity

Among two studies (covering 1997-2003), 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

Fifteen studies (covering 1973-2003) found 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.

Tumor stage

Nineteen studies (covering 1973-2006) 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

Fourteen studies (covering 1973-2006) 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

Seven studies (covering 1989-2006) 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).

Biopsy frequency

Three studies (covering 1982-1996) reported information on trends in the performance of prostate biopsies. The SEER-Detroit study 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. The 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.

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.

Histopathologic grading changes

One study reported on a single pathologist who regraded pathology slides in 2002-04 from patients diagnosed in 1990-92. The regrading resulted in the assignment of significantly higher 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 timepoints. 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 16 studies (covering 1973-2008), most demonstrated increasing trends in the proportion of patients being managed with observational management strategies of no active treatment (AS, WW or expectant management), with or without androgen deprivation therapy (ADT). In all six studies providing data since 2000, the proportion of patients receiving AS or WW was less than 10 percent; this also held true for subgroups of patients with "low-risk disease" investigated in two studies.

Key Question 2

How are active surveillance and other observational management strategies defined?

Because the terms AS and WW (as well as others) have been used by investigators to denote strategies both with or without curative intents, we divided protocols into those which had been clearly described as curative and those in which their aims were either unclear or primarily palliative, regardless of how these regimens were labeled.

Strategies with curative intent

Fifteen 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), there was little common ground regarding eligibility criteria for AS. 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 to men under age 70 or 80 years.

Gleason score. Twelve cohorts based eligibility for AS on a Gleason score. However, there was little agreement across cohorts in the threshold used. The most frequent criteria were Gleason score ≤ 6 (7 studies) and ≤ 3+3 without Gleason 4 or 5 pattern (5 studies).

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

Prostate specific antigen (PSA). Ten of 15 cohorts used PSA as part of the eligibility criteria for AS. Studies used a wide range of thresholds, mostly ranging from 10 to 15 ng/mL. Two cohorts used PSA density (PSA per volume of prostate tissue) thresholds.

Imaging. Nine cohorts required that biopsies be performed with transrectal ultrasonography. One of these cohorts also required a chest radiograph. 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 15 cohorts included regular PSA testing in the followup protocol but there were no uniform monitoring frequencies; 11 cohorts also included regular digital rectal examination (DRE), also at various frequencies; and 13 performed routine rebiopsy between 6 and 48 months. One cohort included a regular bone scan schedule. Criteria for recommending curative treatments varied widely across all 15 AS cohorts. The recommended treatments were also not standardized and were determined by the physicians in many of the cohorts.

Gleason score. Eleven cohorts described using Gleason score as part of their monitoring criteria for disease progression. Generally, progression in Gleason score 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 explicitly 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 (6 cohorts) and greater than four (3 cohorts) positive biopsy cores. 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 monitoring criteria for disease progression but specific percentage cancer involvement was not reported.

Prostate specific antigen. All 15 cohorts included regular PSA testing in the followup protocol but there were no uniform followup frequencies. Five cohorts considered rising PSA and/or PSA kinetics as part of triggers for treatment but did not specify the detailed criteria. Eight cohorts used a variety of PSA triggers for treatment.

Imaging. Six cohorts performed transrectal ultrasonography-guided biopsies. One of these cohorts also performed 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. One cohort reported that some patients requested treatment due to anxiety related to increasing PSA.

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 (PSA). Five cohorts used PSA within 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 norma chest radiograph findings.

Behavioral indications. No cohort explicitly 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 (PSA). 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 15 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 values. AS protocols generally did not include imaging in their followup protocols. In contrast, other observational strategies typically included imaging in their followups, specifically bone scan and chest radiography. They also generally did not use rebiopsy but they did use PSA in their followups. Comparison of the followup frequencies between AS and other observational strategies (Tables 2.4 versus Table 2.8) 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 of, acceptance of, and 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 offer, acceptance, or adherence of AS (however, no such studies were found). 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.

Primary care

One survey of New Zealand general practitioners found that 45 percent would offer observational management strategy if the patient's life expectancy was <10 years, but only 3 percent would offer observational management strategy in patients with 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 strategy 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 – 2nd opinion

One survey of men diagnosed with early-stage cancer who had not yet decided on treatment 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 four 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 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 receiving observational management strategies was predicted by older age, an increased number of comorbidities, lower Gleason score, well-differentiated tumor, lower stage disease, lower PSA, and 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 or increased free-to-total PSA ratio or more rapid PSA increase, but not comorbidities, 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. 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

An analysis comparing men who refused randomization but selected AS to men who were randomized to AS found that lower baseline anxiety was associated with the decision to choose AS (and not be randomized). Three 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 affects 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 was 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 insurance or Veterans Administration (VA) insurance. One analysis reported that having preferred provider organization (PPO) or health maintenance organization (HMO) coverage decreased the probability of receiving observational management strategies versus RP. It also reported that Medicare supplemented with fee-for-service, HMO, or PPO 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 3 major cities). Comparisons among other registries were nonsignificant.

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 (versus 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 versus 44 percent, respectively).

Academic centers versus private practice

One multivariable analysis reported that treatment facility status (academic versus 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.

Decision-making 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?

No study reported clinical outcomes specifically of 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.

Due to the lack of studies comparing AS to immediate treatment, we evaluated studies that compared observational management strategies (largely resembling WW) with immediate treatment. In addition to previously published systematic reviews and evidence reports, our searches identified one multicenter RCT (2 publications, one on clinical outcomes and one on costs) and 12 cohort studies (2 prospective and 10 retrospective). Of note, 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 treatments.

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 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 were heterogeneous.

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 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 were heterogeneous.

Observational management strategies versus combined active treatments or combined radiation treatment modalities

One study reported 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 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 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 suggests that AS may be associated with higher costs compared to RP or BT, but lower costs compared to intensity modulated RT (IMRT) or proton beam RT.

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 or standardizations 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.

Key question 1 - Patient population and natural history changes in last 30 years

Better understanding of time-trends can be gained by improving the data being collected and expanding the scope of the major U.S. databases. In particular, stage and grade information are often incomplete requiring researchers to create broad categories that place major limitations on the analyses. The SEER database is inadequate to analyze data from races other then 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 report using the “best available information” on staging information because the “best available staging information” will depend 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 that is available prior to surgery.

Key question 2 - Definition of active surveillance

Little new research per se 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 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; 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) criteria or triggers for stopping AS to seek definitive treatments.

Assuming that AS is an intervention plan that many patients may elect (if offered) 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 metastatic disease 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 or 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 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 not to follow the protocols.

Key question 3 - Factors that affect offer, acceptance, and adherence to AS

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 the 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 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 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.

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

The least biased, most reliable study design comparing two interventions is the well-conducted randomized controlled trial that adheres to modern standards. Outcome assessors—particularly those who conduct psychometric testing—should be blinded. The primary outcomes of interest should be patient-centered clinical outcomes, including psychometric tests, adverse events, resource utilization, and costs. Trials need to be of sufficiently long duration to collect data on the clinically relevant outcomes.

In lieu of randomized trials, adequate findings may be possible 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 having adequate data for unbiased analyses.

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*. The factors listed in Key Questions 1 and 2 form a good starting point to consider which subgroups may be of interest.

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. 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 for both blacks and whites. Another consistent trend over time 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 5-7). It has been hypothesized that this effect is caused by changes in histopathological grading guidelines, a preference towards avoiding assigning Gleason 2-4 scores based on prostate cancer biopsy samples, 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 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 15 unique cohorts formed in the PSA screening era that used different formal protocols to monitor triggers for curative treatment of prostate cancer. In all, 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 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 the cohorts.

Because of the different usages 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 which patients had true AS or WW, or who were simply not treated (for a variety of reasons), or who had 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 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 an observational management strategy: older age, presence of comorbidities, higher Gleason score, higher tumor stage, higher diagnostic PSA, higher 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, 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 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 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. 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 costs 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 Institute for Clinical and Economic Review (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 IMRT or proton beam RT.

Introduction

In 2011, over 240,000 men are projected to be diagnosed with prostate cancer and 33,000 to die from the condition.[1] Currently, in the United States, most instances 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.[2] 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. Notably, though, 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 much likelihood of obtaining any clinical benefit due to the slow progression of many prostate tumors.[2] However, 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 generally connotes the monitoring of a potentially curable prostate cancer and intervening with a curative-intent treatment at the earliest sign of worrisome progression. In contrast, WW generally connotes postponing therapeutic interventions until symptom development, with the primary objective being palliation of the symptoms rather than an attempt at a cure. 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 a relatively passive strategy—with interventions triggered by symptoms. It should be underscored, however, that in the scientific literature the two terms and their intents are often used interchangeably.

Given the tradeoffs between complications from curative treatments and long-term risks of delaying treatment, and thus the use of AS and other observational management strategies by men who are more interested in avoiding the risks of curative treatment, it is important to clarify appropriate eligibility criteria and followup protocols for the 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. It is also of interest to evaluate whether men offered AS will accept this strategy and adhere to it. If men feel a strong need “to do something” to definitively treat the cancer, and thus AS is rarely chosen or not adhered to, then the impact of offering this strategy will be small. Therefore, factors that relate to the offer of AS by clinicians to patients, acceptance of AS by patients and their families, and adherence with AS once this course has been chosen need exploration.

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: http://consensus.nih.gov/

The Conference planning committee crafted the following key questions related to the natural history of prostate cancer, the definitions of AS, the factors pertained to the practice of AS, the comparative effectiveness of AS, and the future research needs in AS. The exact wording of the Key Questions provided to the EPC to be addressed by systematic review 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. Follow up 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 – 2nd 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?
   f. Prostate specific and all cause mortality
   g. Morbidity of primary treatment decision
   h. Incidence of metastatic disease
   i. Quality of life
   j. 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 U.S. 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*.[3]

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 Preface for the list of members of the TEP, and title page for our local domain experts.

Key Questions

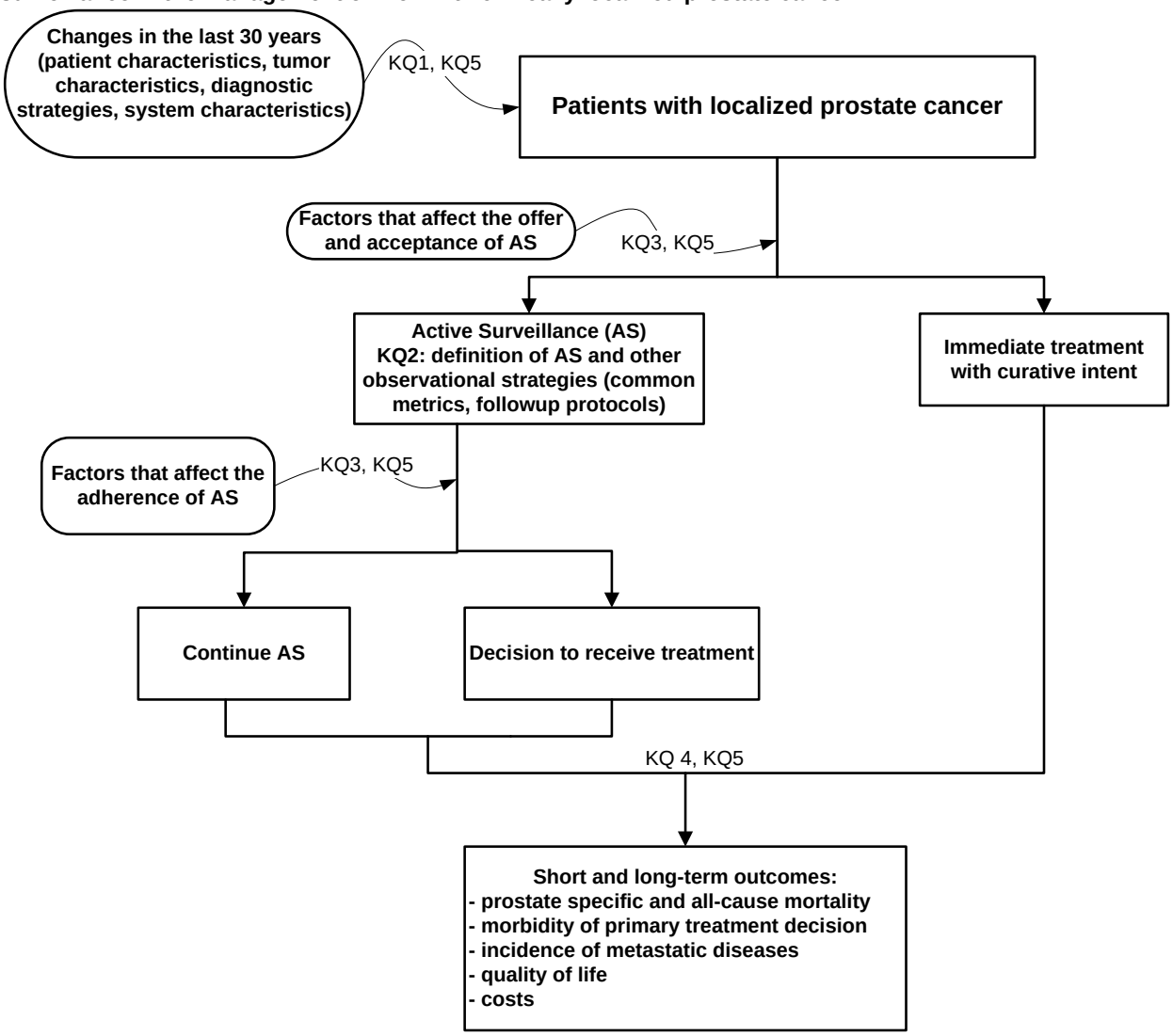
The Key Questions listed in the Introduction were provided by the NIH Consensus Development Program (CDP). The Key Questions have not been altered for the review.

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 (KQ) examining the role of active surveillance in the management of men with clinically localized prostate cancer



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 March 4, 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 April 14, 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.[4]

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 consider 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.

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, decision analyses

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

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.[5,6] For treatment-related costs, we also considered two economic evaluations[7,8] prepared by the Institute for Clinical and Economic Review (ICER) of the Massachusetts General Hospital and an associated journal publication.[9] ICER reports are available online at http://www.icer-review.org (last accessed: August 7th, 2011). We did not consider the treatment effectiveness components of the ICER reports because they did not provide additional 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 registry 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 consider 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 sourced 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 two 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. 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.

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 Prostrate Cancer Outcomes Study [PCOS], or the Patterns of Care Study [POCS]). 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.

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.[3] 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 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 website (http://seer.cancer.gov/; last accessed July 18th, 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 sampling population, 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) or standard deviations (SD) and when available, the mean difference (between groups), 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[3] and an updated 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 high, medium or low 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 U.S. 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.

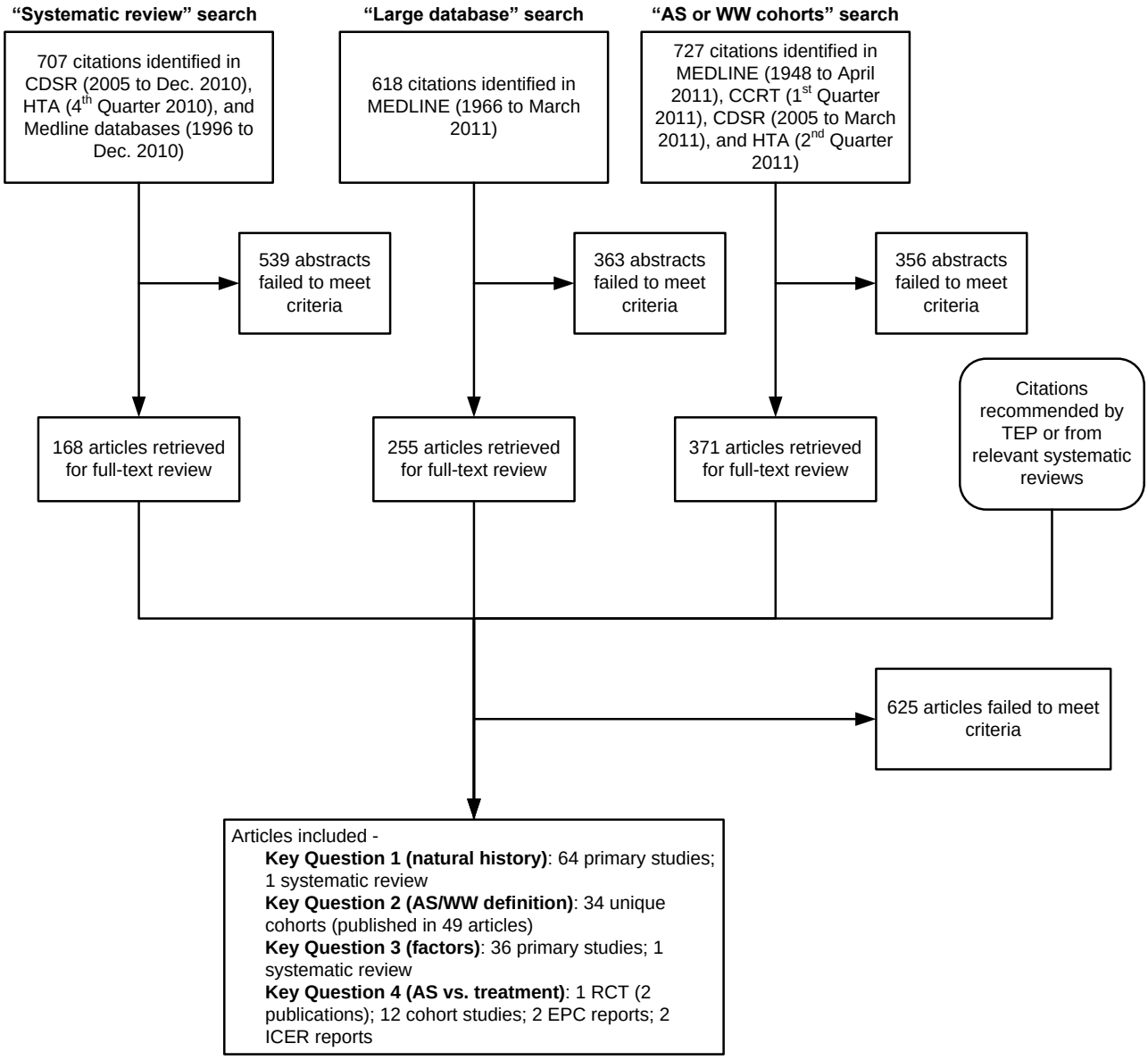
Peer Review and Public Commentary

[Pending]

Results

Our literature searches for prostate cancer systematic reviews, for large registry databases, and for AS/WW publications yielded 701, 618, and 727 citations, respectively. From these, 794 articles were provisionally accepted for review based on the abstracts and titles. Additional citations recommended by technical expert panel or from reference lists of relevant systematic reviews were also accepted for review. After screening their full texts, 625 articles were rejected for not meeting eligibility criteria. In total, 169 articles met criteria and are reviewed.

Figure 2. Literature flow



The numbers of studies for each Key Question do not sum to the total number of studies because some studies addressed multiple Key Questions. AS = active surveillance; CDSR = Cochrane Database of Systematic Reviews; HTA = Health Technology Assessment; CCRT = Cochrane Central Register of Controlled Trials; TEP = technical expert panel; RCTs = randomized controlled trials; EPC = Evidence-based Practice Center; ICER = Institute for Clinical and Economic Reviews; 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 U.S. 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 would provide information stratified by factors relevant to Key Question 1 (see the end of the Introduction for the list of factors).

We identified 64 primary observational studies and two systematic reviews eligible for inclusion in Key Question 1.[2,16-78]

Of the primary observational studies, 41 analyzed the Surveillance Epidemiology and End Results (SEER) database of the National Cancer Institute (NCI) or a subset of its component registries, six additional studies utilized the linked SEER-Medicare database, ten the Cancer of the Prostate Urologic Research Endeavor (CaPSURE) database, four 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.[a]
- 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.[79,80]
- The SEER-Medicare database linked the SEER cancer registries data and Medicare enrollment and claims files.[b]
- A comparison of sociodemographic characteristics of Medicare beneficiaries residing in the SEER areas versus the general U.S. elderly population has demonstrated that the age and sex distribution for individuals 65 years and older in the SEER areas is 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.[81]
- 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] http://seer.cancer.gov/registries/characteristics.html ; last accessed: July 18th, 2011.
[b] See http://healthservices.cancer.gov/seermedicare/; last accessed: July 18, 2011.

collects data from 40 urology practices in the U.S. (34 community based, 3 Veterans Administration [VA])-based and 3 academic center based).[c]

- 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.[d]
- 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.[e]

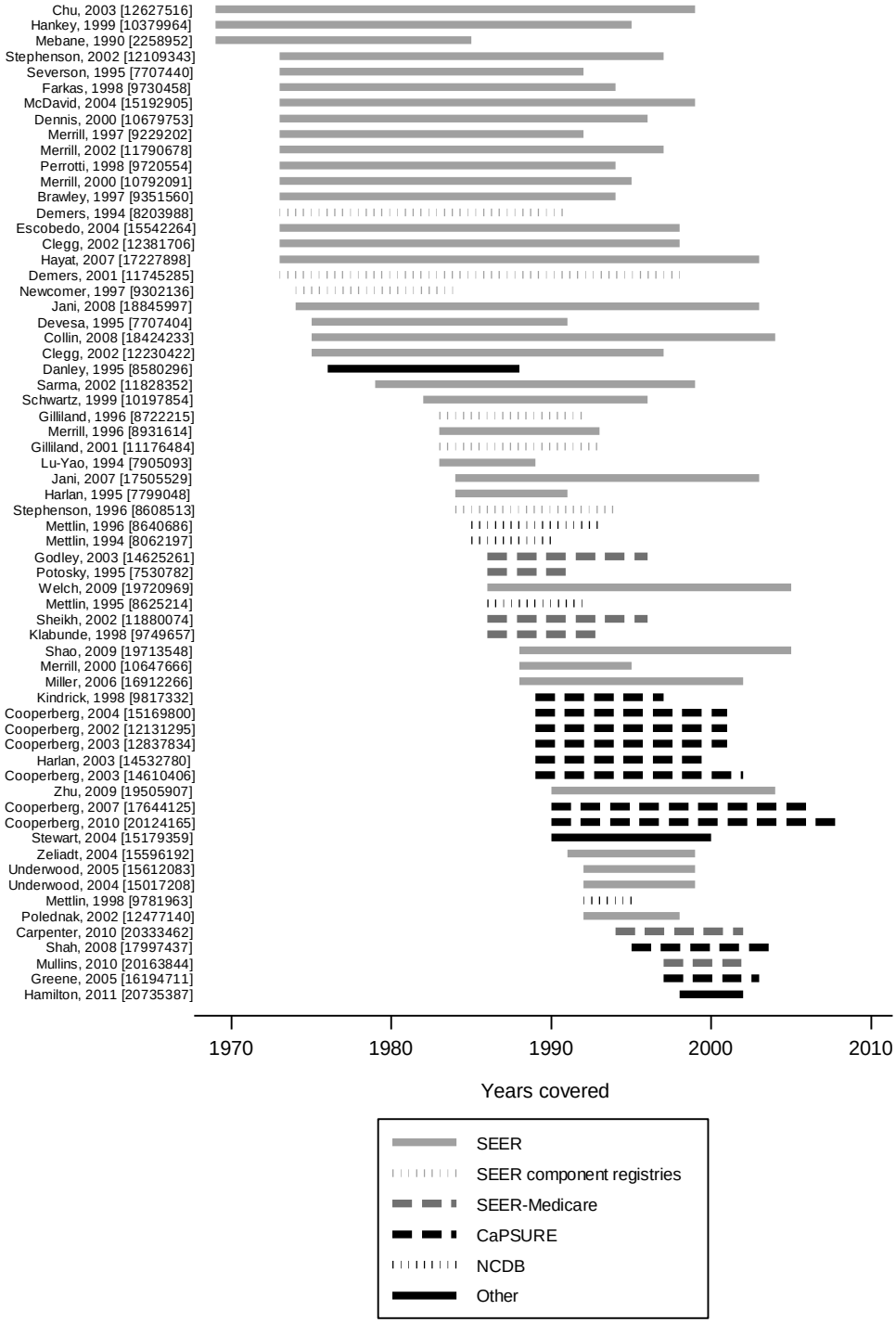
Included studies had large sample sizes (median sample size = 45,119 patients; 25th-75th percentile 9626-138,387), were published between 1990 and 2011, and analyzed data from 1969 to 2008. Figure 1.1 presents the years covered by each primary study and the databases used. Appendix Tables C1.1-C1.12 present additional information about each of the studies relevant to Key Question 1.

We organized the 64 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

[c] See http://urology.ucsf.edu/clinicalres/CRuroOnc_gceps_capsure.html; last accessed: July 18, 2011.
[d] See http://www.facs.org/cancer/ncdb/index.html; last accessed: July 18, 2011.
[e] See http://www.cdc.gov/nchs/; last accessed: July 18, 2011.

Figure 1.1. Years covered and databases utilized by studies considered eligible for Key Question 1

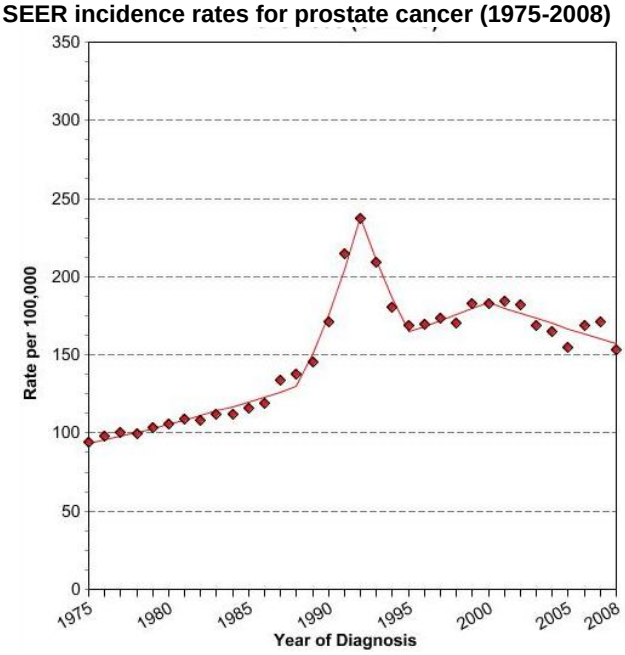


Horizontal lines indicate the years covered by each primary study considered for Key Question 1. Different line patterns indicate the different databases utilized by each study. Studies are listed by the first year covered, then by database used, then by year of publication and are presented using the 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.

1. Trends in prostate cancer incidence

Prostate cancer incidence trends in the U.S. 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 rose between 1975 and 1992 and then fell until around 1995. After a period of nonsignificant increase from 1995 to 2000, rates declined again from 2000 to 2007 (**Figure 1.2**).

Figure 1.2. 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: July 18th, 2011).

Overall, 28 studies provided information on trends in prostate cancer incidence. Of these, 10 provided information stratified by patient age,[17,27,28,35,38,39,43,55,57,59] 14 by race/ethnicity,[16,17,19,21,25,32,34,42,43,45,46,49,57,77] 13 by tumor stage,[17,18,20,23,26,27,34,35,42,50,52,55,56] and 6 by tumor grade.[17,21,42,52,56,77] 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 = 46,248; 25th-75th percentile 33,086-156,598), were published between 1990 and 2009, and provided information for years 1969 to 2005.

Patient Characteristics

Age

Ten studies (8 SEER, 2 SEER-Medicare) covering 1973 to 2005 provided information on prostate cancer incidence stratified by patient age.[17,27,28,35,38,39,43,55,57,59] 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

Fourteen studies (12 SEER, 1 SEER-Medicare, 1 LAC/USC) covering 1973 to 2005 provided information on prostate cancer incidence stratified by patient race/ethnicity.[16,17,19,21,25,32,34,42,43,45,46,49,57,77] Twelve of the 14 studies provided information exclusively for whites or blacks and only two provided information on patients belonging to other racial or ethnic groups; data for other racial/ethnic groups were 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 provided information up to 2005, demonstrating 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

Thirteen studies (12 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] Twelve studies investigated trends in the incidence of localized/regional and distant disease. Generally, 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. 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

Six 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).[17,21,42,52,56,77] 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]

2. 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]

Among the studies we reviewed, 17 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. Eight of the studies provided information stratified by age,[17,18,24,28,30,33,39,40] 13 by race/ethnicity,[16-19,24,26,27,34,40,44,51,60,77] one by comorbidity status,[24] three by tumor stage,[18,24,40] and two by tumor grade.[24,40] No studies provided information stratified by any of the other factors relevant to Key Question 1. Of the studies considered eligible, 14 utilized the SEER database, one the SEER-Medicare database and two other databases. Studies were generally large (median sample = 60,494; 25th-75th percentile 42,269-245,510), were published between 1990 and 2008, and covered years 1969 to 2003.

Patient Characteristics

Age

Eight studies (all SEER-based) covering 1969 to 2003, reported information on population mortality rates (4 studies) or prostate cancer-specific survival among patients diagnosed with prostate cancer survival (3 studies), stratified by age.[17,18,24,28,30,33,39,40] Studies of population mortality rates demonstrated decreases in the mortality rate for all age groups between the early-1990s and 1999. No study reported information for years after 2000.

One study of prostate cancer-specific survival, using the SEER database, demonstrated that over time (for death 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) prostate cancer patients with multiple primary cancers were consistently less likely to die of prostate

[a] http://progressreport.cancer.gov/

cancer compared to patients with no multiple primary cancers. No studies provided information on comorbidities other than multiple primary cancers.

Race/ethnicity

Thirteen studies (10 SEER or component registries, 1 SEER-Medicare, 1 LAC/USC and one using data from the NCHS) covered 1969 to 2000 and reported information on trends in prostate cancer mortality rates (9 studies) or prostate cancer-specific survival (6 studies; 2 studies reported both types of information) stratified by patient race/ethnicity.[16-19,24,26,27,34,40,44,51,60,77] Ten of the studies reported exclusively on non-Hispanic whites or blacks whereas three 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 2000 for all racial/ethnic groups. No study provided information on mortality rates for the years after 2000. 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. Five of the six 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 non-prostate 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

Three studies (all SEER based) reported on temporal trends in prostate cancer specific mortality among patients diagnosed with prostate cancer.[18,24,40] Studies covered 1969 to 1999 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

Two studies (both SEER based) including prostate cancers diagnosed between 1973 to 1995, reported information on trends in prostate cancer survival stratified by tumor grade at diagnosis.[24,40] 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]

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

3. Patient, tumor, and system-level characteristics at diagnosis

Patient Characteristics

We identified 43 observational studies reporting on patient characteristics at presentation (24 SEER, 4 SEER-Medicare, 9 CaPSURE, 4 NCDB and 2 other databases).

The most commonly examined characteristics at baseline were patient age (19 studies),[19,21,29,30,37,42,46,50,52,55,56,63,65,72-76,78] race/ethnicity (15 studies),[31,37,46,49,52,53,58,61-63,65,72,75,76,78] tumor grade (14 studies),[2,22,30,37,46,52,56,65,68,69,72,74,76,78] and tumor stage (19 studies).[2,16,23,31,37,46,52,56,62,63,65,68,69,72-77] Information was available for all factors relevant to Key Question 1 except tumor volume. Studies were generally large (median sample = 41,433; 25th-75th percentile, 8215-138,387), published between 1990 and 2011, and covered the years 1973 to 2008.

Age

Nineteen studies (11 SEER or its component registries, 1 SEER-Medicare, 2 CaPSURE, 4 NCDB, 1 Patterns of Care Study (POCS) data) covering 1973 to 2005, reported information regarding patients' age at presentation.[19,21,29,30,37,42,46,50,52,55,56,63,65,72-76,78]

Six studies (4 SEER or its component registries, 2 NCDB) reported trends in the average (mean or median) age at diagnosis of prostate cancer. Four 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 two studies (both 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);[21,42] no analyses were reported for individuals belonging to other racial/ethnic groups.

The remaining 12 studies (7 SEER or its component registries, 1 SEER-Medicare, 2 CaPSURE, 1 NCDB, and 1 POCS) reported the distribution of patients across discrete age groups 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.[72,78] 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.[72] 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).[78] 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

Fifteen studies (7 SEER, 3 SEER-Medicare, 2 CaPSURE, 2 NCDB, 1 POCS) covering 1973 to 2003, reported information on trends in the racial/ethnic distribution of patients with prostate cancer.[31,37,46,49,52,53,58,61-63,65,72,75,76,78] Five of the studies analyzed only whites and blacks; the remaining 10 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.

Tumor Characteristics

Stage

Nineteen studies (7 SEER, 2 SEER-Medicare, 5 CaPSURE, 4 NCDB, 1 LAC/USC) covering 1973 to 2006, reported information on trends in the distribution of prostate cancer stage at diagnosis. [2,16,23,31,37,46,52,56,62,63,65,68,69,72-77]Thirteen 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). Five 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 and the single study reporting on T1-T3 tumors 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). 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.

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

Fourteen studies (6 SEER or its component registries, 5 CaPSURE, 2 NCDB, 1 POCS) covering 1973 to 2006 reported information on trends in tumor grade distribution at disease presentation.[2,22,30,37,46,52,56,65,68,69,72,74,76,78] All studies 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 all ten studies that reported the results of statistical tests assessing changes in the distribution of tumor grade over time.

PSA

Seven studies (6 CaPSURE, 1 POCS), covering 1989 to 2006, reported information on trends in PSA levels at presentation.[2,64,65,68,69,72,78] Six 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 levels below 10 ng/mL).

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).[82] 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 test (e.g., as part of the investigation of clinical symptoms suggestive of prostate cancer) is particularly challenging using administrative data and may be uncertain even after review of complete medical records.[82]

Diagnostic Strategies

Biopsy Frequency

Three studies (2 SEER component registries, 1 SEER-Medicare), covering 1982 to 1996, reported information on trends in the performance of prostate biopsies.[52,54,57] 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] The 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]

Number of cores

A single primary study[71] provided information on the number of biopsy cores obtained during the investigation of suspected prostate cancer cases prostate cancer. The study utilized the CaPSURE database and covered 1997 to 2002. It demonstrated a significant increase in the mean number of cores examined (from 7.5 in 1997 to 9.8 in 2002) per patient. 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.[83] 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.[84]

We also identified a structured review on the same topic through additional targeted searches[85]. 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.[65,72,75,78] Three studies (2 CaPSURE, 1 POCS) provided information on trends in the distribution of patients' insurance status at diagnosis.[65,72,78] 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).[65,72] In contrast, the study using POCS data did not demonstrate a change in the distribution of insurance status over time (1998-2002).[78]

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 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.[78]

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.[75]

4. Trends in treatment patterns

Sixteen studies (5 SEER, 1 SEER-Medicare, 5 CaPSURE, 4 NCDB, 1 POCS) provided information on treatment trends over time.[2,27,36,41,47,48,58,67,68,70,72-76,78] Studies were generally large

[b] The Will-Rogers phenomenon arises whenever 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.

(median sample size = 60,304; 25th-75th percentile 5828-125,529), published between 1994 and 2011, and covered 1973 to 2008. In eight 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). 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 six studies (5 using 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/Conclusions

We reviewed 64 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 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 (Gleason score 2-4 and >7, respectively) tumors, and a concomitant increase in intermediate grade tumors (Gleason 5-7). Over time, patients diagnosed with prostate cancer are less likely to die of the disease (i.e., they are more likely to die of non-prostate 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.

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 treatments 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 typically uses a predefined protocol to monitor triggers for curative treatment of prostate cancer, and watchful waiting (WW) uses a somewhat passive (compared to AS) followup and palliative or (potentially) curative treatments are instituted when the patients become symptomatic. 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.[4]

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 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 Tables C2.1, C2.2). 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.

Because the terms AS and WW (as well as others) have been used by investigators to denote strategies both with or without curative intents, we divided protocols into those which had been clearly described as curative and those in which their aims were either unclear or primarily palliative, regardless of how these regimens were labeled. In the following sections, we first describe published protocols for the cohorts with clearly reported curative intent, followed by a description of the protocols for observational management strategies with primarily palliative or unclear treatment intent.

Protocols with curative intent

We identified 15 unique cohorts reporting formal protocols to monitor triggers for curative treatment of prostate cancer. The triggers for curative treatment had to include parameters other than symptomatic disease progression. Of these cohorts, seven are in the U.S., two in Canada, two in the UK, one in the Netherlands, and one in Japan (Table 2.1). 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 15 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 2.1. Unique AS cohorts

Country	Center or Study Name
United States	Baylor college of Medicine Cleveland Clinic Foundation Dana-Farber Cancer Institute John Hopkins University Memorial Sloan-Kettering Cancer Center (MSKCC) University of California at San Francisco (UCSF) University of Connecticut Health Center University of British Columbia University of Miami
Canada	McGill University Toronto-Sunnybrook Regional Cancer Center
United Kingdom	Royal Marsden Hospital (since 1993)* ProtecT (Prostate testing for cancer and Treatment) trial
Netherlands	PRIAS (Prostate cancer Research International Active Surveillance) study
Japan	Kagawa Medical University (since 2002)**

*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.
** 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.

a. Common metrics: eligibility criteria for low-risk or clinically localized prostate cancer in AS cohorts (Table 2.2)

There are no uniform criteria used to identify patients with low-risk or clinically localized prostate cancer across these 15 cohorts, with the exception of 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). Nine cohorts explicitly reported that transrectal ultrasound (TRUS) guided biopsy was used to confirm the diagnosis of prostate cancer.

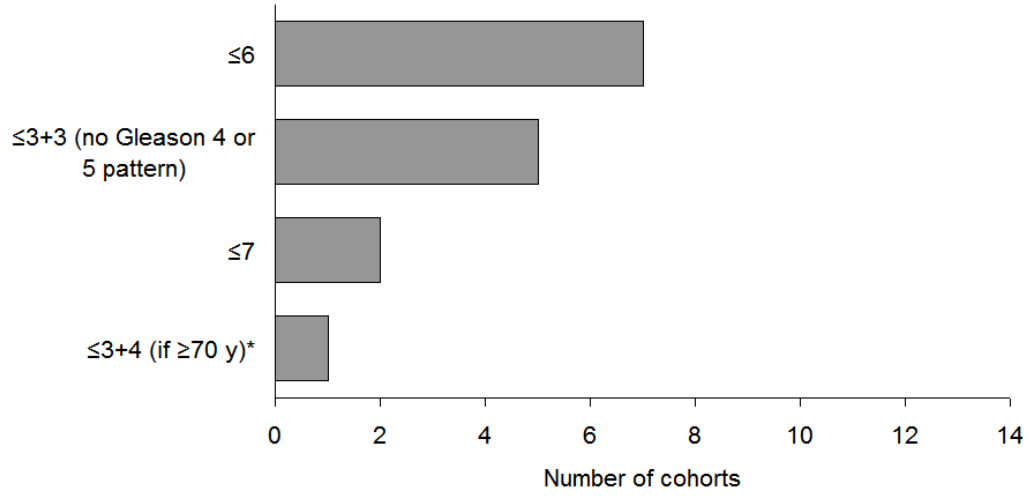
i. Age

Three cohorts reported age as part of patient eligibility criteria for AS.[86-88] 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),[86] less than 80 years in the cohort at University of Miami,[87] and between 50 and 80 years old in the cohort at Kagawa Medical University in Japan.[88] Only 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 diagnossis;[86] the other two did not report the reason for including age as part of patient eligibility criteria for AS.[87,88]

ii. Gleason score

Twelve cohorts (16 publications) used Gleason score as part of patient eligibility criteria for AS (Figure 2.1).[86-101] Of these, 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.[102] 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.[95,96]

Figure 2.1. Summary of 12 cohorts that used Gleason score as part of AS program eligibility criteria



Legends: The numbers of cohorts for each criterion do not sum to the total number of cohorts because some cohorts used multiple criteria. 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).
* ≤3+4 (if ≥70 yr) = for older patients (age ≥70 years), less than Gleason up to 3+4 criterion was used.

iii. Number of cores positive for cancer

Eight cohorts (11 publications) used maximal number of biopsy cores positive for cancer as part of patient eligibility criteria for AS.[86-88,90-94,97,98,100] 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).

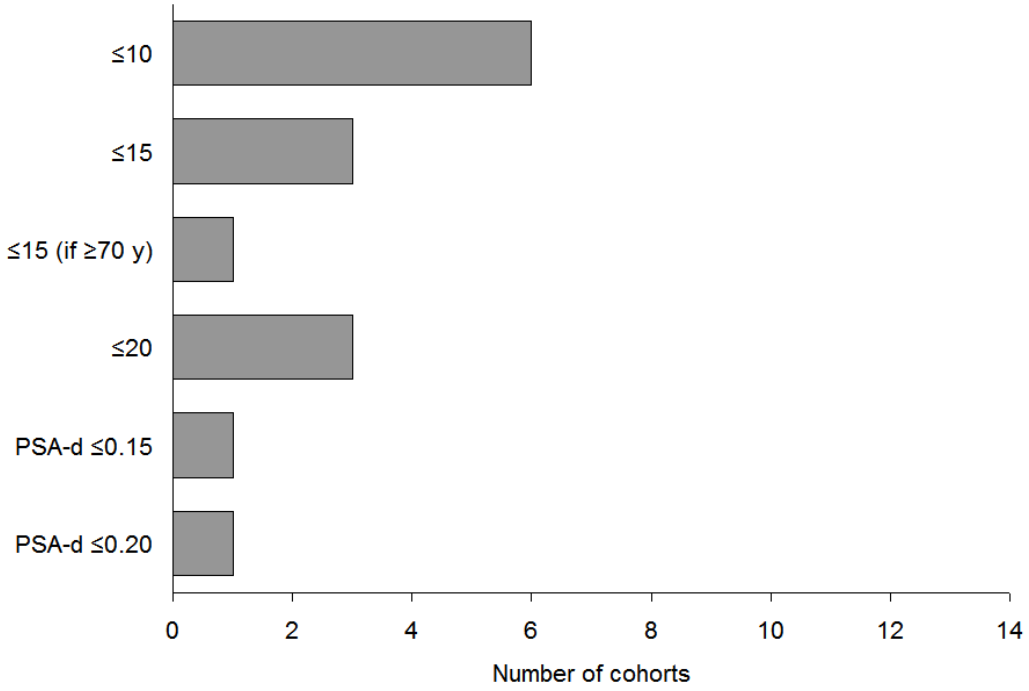
iv. Percentage cancer involvement in each core

Five cohorts (7 publications) used "low-volume disease" as part of patient eligibility criteria for AS.[87,90,91,94,97,98,101] In three cohorts, the definition of "low-volume disease" was less than half of the cancer involvement in any individual core. In the other two cohorts (3 publications), the criterion was described variably as less than half of two biopsy cores,[90] less than 20 percent in one or two biopsy cores,[91] and cancer involvement of less than 33 percent of biopsy cores.[101]

v. PSA

Ten cohorts (14 publications) used PSA as part of patient eligibility criteria for AS (Figure 2.2).[86-88,90-97,99-101] 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,[92,93] 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.[87,90,91]

Figure 2.2. Summary of 10 cohorts that used PSA (ng/mL) as part of AS program eligibility criteria



Legends: The numbers of cohorts for each criterion do not sum to the total number of cohorts because some cohorts used multiple criteria. PSA-d = PSA density (ng/mL/cm^3) which is calculated by dividing the PSA level by the total volume of the prostate.

vi. Imaging

Nine cohorts (12 publications) performed TRUS guided biopsy in confirming the diagnosis of prostate cancer.[87-92,94-96,100,101,103] Of these, one cohort also 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.[95,96] One other cohort reported that magnetic resonance imaging of the prostate was selectively used at diagnosis.[86]

vii. Behavioral indicators

No behavioral indicator was used explicitly as a criterion for AS program enrollment. Only one cohort surveyed patients' reasons for choosing AS, and they found that physician influence had the greatest impact on choosing AS.[87] Other behavioral factors for choosing AS included concerns for incontinence and erectile dysfunction.[87] Another cohort reported that some patients enrolled in AS had limited life expectancy due to advanced age or poor medical condition.[103]

Table 2.2. 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[89] [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[103] [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[104] [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,[i] US[86] [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[87,90,91] [17850361; 20800964; 21215429] 1991-2007	AS; WW[91]	≤80[87]	≤6[90]	≤50% of 2 biopsy cores[90] ≤2 biopsy cores with ≤20% in each core[91]	≤15[90] ≤10[87,91]	TRUS guided biopsy	≤T2/T2b[90]	MD influence had the greatest impact on choosing AS (73%), concerns for incontinence (48%) and erectile dysfunction (44%) also reasons for choosing AS[87]
UCSF, US[101] [18433013] >1991	AS	–	<10	≤6; absence of Gleason grade 4 or 5	cancer involvement of <33% of biopsy cores	TRUS guided biopsy every 6-12 mo	T1/T2a	–

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)
Royal Marsden Hospital, UK[92,93] [15839912; 17850368] 1993-2002; ≥2002[93]	AS	–	≤7 (primary ≤3)	Less than half of the biopsy cores positive (octant biopsy).[93]	≤20[92] <15[93]	TRUS guided biopsy	T1–2, N0/X, M0/X	–
John Hopkins, US[94] [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[95,96] [11395227; 19917860] 1995-2002 as a phase II trial; 2003-ongoing as an open prospective cohort	WW; AS[ii]	–	≤7[95] ≤6; ≤3+4 (if ≥70 yr)[96]	–	≤15[95] ≤10; <15 (if ≥70 yr)[96]	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)[iii]	–
Memorial Sloan-Kettering Cancer Center, US[97] [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[105] [19603015] 2000-2008	Active monitoring	–	–	–	–	–	"Clinically localized prostate cancer"	–
Dana-Farber Cancer Institute, US[98] [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[88] [18272471] 2002-2003	AS	50-80	≤6	1-2 positive cores per 6-12 systematic biopsy cores	≤20	TRUS-guided six sextant biopsy	T1cN0M0	–

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)
Cleveland clinic, US[99] [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[100] [19817747] 2006 – ongoing	AS	–	≤3+3=6	Adequate biopsy sampling according to biopsy protocol; maximal 2 biopsy cores invaded with prostate cancer	≤10 PSA density ≤ 0.2 ng/ml/ml	TRUS guided biopsy[iv]	T1c or T2	–

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

[i] Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami
[ii] Multiple terms used across multiple publications and within single publication
[iii] Since January 2000, Toronto-SRCC study was restricted to low-risk patients only.
[iv] PRIAS protocol can be found: http://www.erspc-media.org/media/publications/PRIAS%20Project_background.pdf (assessed 7/15/2011)

Table 2.3. 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[89] [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	TRUS guided biopsy	–	–	Definitive treatment when objective progression or patients' requests.
McGill Univ., Canada[26,103] [18484590] 1987-2002	Every 3-6 mo PSA and DRE	Gleason pattern of 4	≥3 positive, or >50% cancer in at least 1 core	√	TRUS guided biopsy was done annually or when there was a change in DRE or PSA.	√	–	Clinical disease progression on DRE or repeated sextant biopsy, patient preference, or rising PSA level.[26]
Univ. of Connecticut Health Center, US[104] [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).
Four tertiary care academic medical canters,[v] US[86] [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[101,106,107] [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)[106]	≥33% of cores or >50% of any core[107]	PSA velocity >0.75 ng/mL/yr	TRUS	–	–	Disease progression; no specific protocol for intervention (implied)

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Univ. of Miami, US[90,108] [17850361; 10759669] 1991-2007	Every 3-4 mo PSA and DRE for 2 yr and every 6 mo thereafter.[90]	≥7[90]	>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.) [90]	Biochemical progression: PSA increase 25-50 %/yr[108]	TRUS (needed for determining tumor volume) [90]	–	–	Treatment is encouraged at an increase in tumor volume, Gleason score ≥7, or the presence of >2 positive cores at rebiopsy.[90] Treatments were offered at the time of local stage progression by DRE and/or biochemical progression, or systemic progression.[108]
Royal Marsden Hospital, UK[92,93,109] [15839912; 17850368; 18949747] 1993-2002; ≥2002[93]	Every 3-6 mo PSA and DRE for 2 yr and every 6 mo thereafter. Rebiopsy not routine.[92] After 2002; monthly PSA in yr 1, every 3 mo in yr 2, and every 6 mo thereafter. DRE every 3 mo for 2 yr.	>7[93] Primary Gleason ≥4, (initial Gleason 3+3, upgraded to Gleason ≥3+4)[110]	TRUS-guided octant biopsy at 18-24 mo. Sextant or octant ≥50% biopsy cores positive.[93]	PSA DT<4 yr[93] PSA velocity >1 ng/mL/yr [109]	TRUS-guided biopsy. Repeat imaging only if clinically indicated.[92]	–	–	Rate of rise of PSA, according to judgment of each patient and clinician.[92] PSA DT<4 yr, histologic progression, or patient preference, or PSA velocity >1 ng/mL/yr [109]
John Hopkins, US[94] [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[111]	–	Annual surveillance biopsy: Gleason ≥7; or Gleason pattern 4 or 5; or >2 cores cancer positive; or single core >50% cancer.

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[95,96,112,113] [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.[96] Sextant biopsies were used from 1995 to 2000; since 2000, 10 to 14-core biopsies were performed using the Vienna nomogram.[112]	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.[vi] In 2005 the group developed a general linear mixed model as a clinical decision making aid.[vii113]	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,[viii] histologic or PSA progression triggered the offer of treatment based on age, extent of disease and comorbidities.
Memorial Sloan-Kettering Cancer Center, US[97] [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.
ProtecT, UK[105] [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"[ix] (implied using PSA level or change and/or rebiospy results as triggers)
Dana-Farber Cancer Institute, US[98] [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[x] were offered surgery or radiotherapy.

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	Gleason score	# biopsy cores /% cores	PSA	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Kagawa Medical Univ., Japan[88] [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[99] [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)
PRIAS, Netherlands[100] [19817747] 2006 – ongoing	PSA at 3 mo, DRE at 6 mo and standard rebiopsy after 1 yr	>3+3=6	Biopsy protocol[xi]	PSA DT 0 to 3 yr	–	–	–	PSA DT 0 to 3 yr, T state >2 or rebiopsy findings exceed study inclusion thresholds

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

---

[v] Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami
[vi] 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.
[vii] 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.
[viii] 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.
[ix] Source: http://www.epi.bris.ac.uk/protect/
[x] 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.
[xi] PRIAS protocol can be found: http://www.erspc-media.org/media/publications/PRIAS%20Project_background.pdf (assessed 7/15/2011)

b. Followup protocols of AS cohorts (Table 2.4)

All 15 cohorts included regular PSA testing in the followup protocol but there were no uniform monitoring frequencies (Table 2.4). DRE was included in the followup protocols of 11 of the cohorts. Thirteen cohorts included routine rebiopsy. One cohort included a regular bone scan schedule. Criteria for recommending curative treatments varied widely across the cohorts. The recommended treatments were also not standardized and were determined by the physicians in many of the cohorts.

i. Gleason score

Eleven cohorts (20 publications) used Gleason score as part of monitoring criteria for disease progression.[89,90,92-101,103,104,106-109,112,113] Generally, progression in Gleason was defined as a Gleason score or pattern greater than those used in the eligibility criteria for AS (Figure 2.1).

ii. Number of cores positive for cancer

Eight cohorts (9 publications) used minimal number of biopsy cores positive for cancer as part of monitoring criteria for disease progression.[88-90,94,97,98,100,103,104] 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.[104] It should be noted that the rebiopsy frequencies varied across the cohorts (Table 2.4).

iii. 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.[92-94,97,98,103,107,109] 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.[99,104]

Table 2.4. Followup frequencies of 15 unique AS cohorts[a]

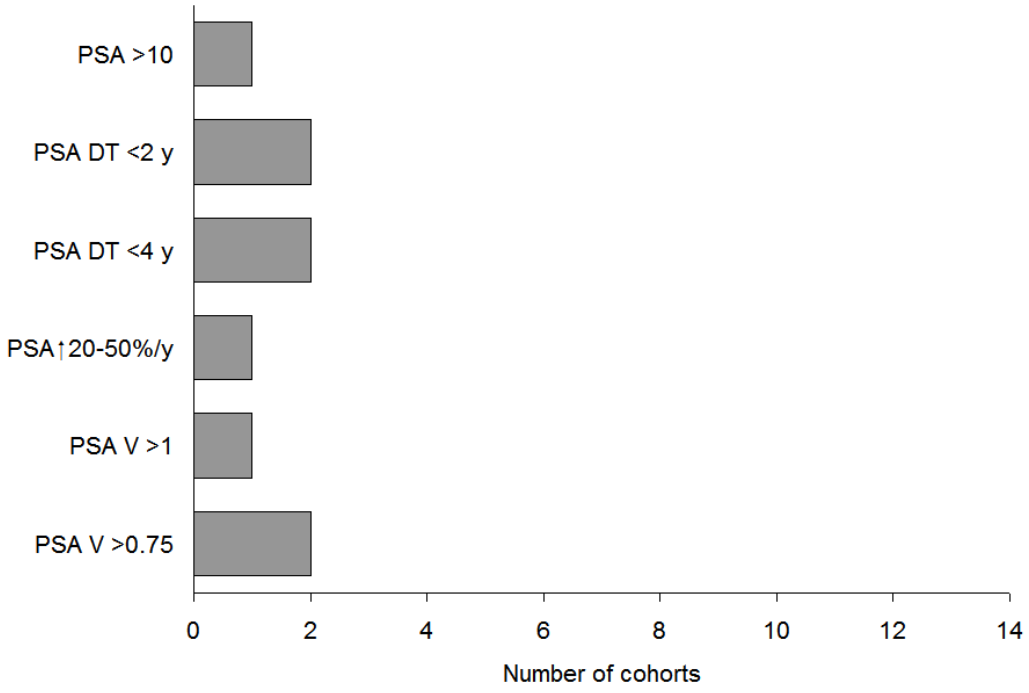
	Year	1				2				3				4			
AS cohort	**Month**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27**	**30**	**33**	**36**	**39**	**42**	**45**	**48**
Cohort 1[89]	PSA-test	√	√	√	√		√		√		√		√		√		√
	DRE	√	√	√	√		√		√		√		√		√		√
	Rebiopsy		√														
Cohort 2[103]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy				√				√				√				√
Cohort 3[104]	PSA-test		√		√		√		√		√		√		√		√
	DRE			√				√				√				√	
	Rebiopsy								√								
Cohort 4[86]	PSA-test			√				√				√				√	
	DRE			√				√				√				√	
	Rebiopsy	within 18 mo; then every 1 to 3 yr															
Cohort 5[90]	PSA-test	√*	√*	√*	√*	√*	√*	√*	√*		√		√		√		√
	DRE	√*	√*	√*	√*	√*	√*	√*	√*		√		√		√		√
	Rebiopsy	within 9-12 mo							√				√				√
Cohort 6[101]	PSA-test	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Rebiopsy	√								√							
Cohort 7[92]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy	not routine															
Cohort 8[94]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy				√				√				√				√
Cohort 9[95]	PSA-test	√	√	√	√	√	√	√	√		√		√		√		√
	DRE																
	Rebiopsy	√**															
	Bone scan				√				√								√
Cohort 10[97]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy	within 12-18 mo; then every 2 to 3 yr															
Cohort 11[105]	PSA-test	√	√	√	√		√		√		√		√		√		√
	DRE																
	Rebiopsy	not routine															
Cohort 12[98]	PSA-test		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√
	Rebiopsy					√				√				√			
Cohort 13[88]	PSA-test	√***		√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE																
	Rebiopsy				√												
Cohort 14[99]	PSA-test			√				√				√				√	
	DRE																
	Rebiopsy								√								√
Cohort 15[100]	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.
Cohort 1 = Baylor College of Medicine and MSKCC; Cohort 2 = McGill University; Cohort 3 = University of Connecticut Health Center; Cohort 4 = Four academic medical centers (Cleveland Clinic Foundation, Memorial Sloan-Kettering Cancer Center, University of British Columbia and University of Miami), Cohort 5 = University of Miami, Cohort 6 = UCSF, Cohort 7 = University of Miami, Cohort 8 = Royal Marsden Hospital, Cohort 9 = John Hopkins, Cohort 10 = Toronto-SRCC, Cohort 11 = MSKCC, Cohort 12 = ProtecT, Cohort 13 = Dana-Farber Cancer Institute, Cohort 14 = Kagawa Medical University, Cohort 15 = Cleveland Clinic
[a] Cohorts of active surveillance in chronological order of starting enrollment year. Some cohorts may include overlapping patients. See Table 2.4 for more detailed monitoring criteria in each cohort.

vi. PSA

All 15 cohorts included regular PSA testing in the followup protocol but there were no uniform followup frequencies (Table 2.4). Five cohorts considered rising PSA and/or PSA kinetics as part of triggers for treatment but did not specify the detailed criteria. Eight cohorts used a variety of PSA triggers for treatment (Figure 2.3). 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.[113] 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).[113] 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.[94,98]

Figure 2.3. Summary of 7 cohorts that used PSA (ng/mL) or PSA kinetics as part of trigger for curative treatment



Legends: The numbers of cohorts for each criterion do not sum to the total number of cohorts because some cohorts used multiple criteria. 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.

v. Imaging

Six cohorts (13 publications) performed TRUS guided rebiopsy,[89,90,92,93,95,96,101,103,106-108,112,113] but there was no uniform rebiopsy frequencies (Table 2.4). Of these, one cohort also performed annual bone scan for the first 2 years and biennially thereafter.[95,96,112,113] Another cohort reported that MRI of the prostate was selectively performed every 1 to 3 years during followup.[86]

vi. Behavioral indicators

One cohort reported that some patients requested treatment due to anxiety related to increasing PSA.[104]

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 U.S.; 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 2.5). 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 2.5. Unique 13 cohorts of observational management strategies with palliative intent

Country	Center or Study Name
US	Howard University college of Medicine* Prostate Cancer Intervention Versus Observation Trial (PIVOT) University of Florida Watchful Waiting Study
UK	Hospitals in Manchester region (University Hospital of South Manchester, Withington Hospital, Christie Hospital; Hope Hospital) Freeman Hospital* Royal Marsden Hospital (before 2002) Western General Hospital*
Sweden	Northern Stockholm region* Orebro Medical Center*
Finland, Sweden, and Iceland	Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4) trial
Netherlands	Erasmus University Hospital
Taiwan	Taichung Veterans Hospital*

*Cohorts that were formed during the pre-PSA screening era.

a. Common metrics: eligibility criteria of observational management strategies with palliative intent (Table 2.6)

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,[114-117] one cohort enrolled "patients without symptoms after initial outflow tract surgery or biopsy,"[118] 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.[119]

Of the seven cohorts in the PSA screening era, three cohorts (4 publications) enrolled patients with clinical stage T2 or less,[111,120-122] one cohort enrolled patients with "low-stage, low-grade disease,"[123] one cohort enrolled patients with any T stage N0/X, M0/X,[92] 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.

i. Age

Four cohorts (5 publications) reported age as part of patient eligibility criteria.[120-122,124,125] The age criterion was less than 75 years in the WW arm in SPCG-4 (Scandinavian Prostate Cancer Group Study Number 4)[124] and PIVOT (Prostate Cancer Intervention Versus Observation Trial),[120] and less than 85 years in the Watchful Waiting Study.[125] One cohort at Erasmus University hospital enrolled patients from the European Randomized Study of Screening for Prostate Cancer (ERSPC).[121,122] 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.[126]

ii. Gleason score

Four cohorts (5 publications) reported Gleason score as part of patient eligibility criteria.[92,121,122,124,125] The Gleason score criterion was less than 8 in three cohorts.[92,121,122,125] 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."[124]

iii. 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.

iv. 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.

v. PSA

Five cohorts (6 publications) used PSA threshold as part of patient eligibility criteria.[120-122,124,125,127] Two criteria were used: PSA less than 50 ng/mL (4 cohorts)[120,124,125,127] and less than or equal to 15 ng/mL (1 cohort).[121,122]

vi. Imaging

Four cohorts (5 publications) used normal bone scan findings as part of patient eligibility criteria.[120-122,124,127] One cohort also required patients to have normal chest radiograph findings to be eligible for the observational management program.[121,122]

vii. Behavioral indicators

No behavioral indicator was used explicitly as a criterion for patient enrollment in the 13 unique cohorts. Two cohorts reported that patients who were enrolled in WW were typically

unsuitable for RP because of advanced age or comorbidities,[92] or had severe medical conditions with a life expectancy of less than 10 years.[123] One cohort reported that "the decision not to treat at diagnosis 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."[121]

b. Followup protocols of observational management strategies with palliative intent (Table 2.7)

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.[114-116,118,119] 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.[117] No information regarding the followup protocol in the pre-PSA screening era was provided. It should be noted monitoring frequencies varied across these six cohorts (Table 2.8).

All seven cohorts (8 publications) in the PSA screening era included regular PSA testing.[92,120-125,127] Again, monitoring frequencies varied across cohorts (Table 2.8). Compared with AS cohorts (see previous section), rebiopsy was not commonly included in the followup protocol among WW cohorts.

i. Gleason score

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

ii. 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.

iii. 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.

vi. PSA

Three cohorts formed in the pre-PSA screening era reported that PSA testing became part of followup protocol after PSA became available.[114,117,125] 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.[92,120,123-125] The sixth cohort reported that "hormonal manipulation was demanded by the protocol when the PSA rose to 50 ng/mL."[127]

v. Imaging

Five cohorts in the pre-PSA screening era included regular bone scan in the followup protocol.[114-116,118,119] Monitoring frequencies varied across cohorts (Table 2.8). One cohort also included regular chest and skeletal radiographs in the followup protocol.[116] Another cohort reported that computed tomography of the pelvis was conducted infrequently.[114]

Three cohorts (4 publications) in the PSA screening era included regular bone scans and chest radiographs in the followup protocol.[120-122,124] Monitoring frequencies varied across cohorts (Table 2.8). Another cohort reported that all patients underwent "multiple bone scans" during followup,[127] so it is unclear whether bone scan was part of followup protocol or not.

vi. Behavioral indicators

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

Table 2.6. 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[114] [1600492] 1967-1989	EM /WW	–	–	–	–	–	Stage A and B	–
Orebro Medical Center, Sweden[115] [7933233] 1977-1984	Deferred treatment	Any (all patients >75 were not given any initial treatment from 1978-79)	–	–	–	–	T0-T2	–
Northern Stockholm region, Sweden[119] [17467883] 1978-1982	WW	–	–	–	–	Normal bone scan	71% T1-2 and 29% with T3	–
Freeman Hospital, UK[118] [3191340] 1978-1985	Deferred treatment	–	–	–	–	–	–	Patients without symptoms after initial outflow tract surgery or biopsy
Western General Hospital, UK[116] [8343901] 1978-1990	Deferred treatment	–	–	–	–	–	incidental (T0/stage A) or localized *T1/stage B1/B2)	–
Taichung Veterans Hospital, Taiwan[117] [12854876] 1983-1996	No treatment	–	–	–	–	–	T1a	–
SPCG-4, Finland, Sweden, and Iceland[124] [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,	No treatment	50-75[xii126]	≤15[122]	<8[122]	–	Bone scan and chest x-	T1c or T2[122]	No treatment decision was made by the urologist in discussion with the patient and his

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)
Netherlands[121] [7544841] ≤1990; 1993-2006[122]	/AS[122]					ray negative		family, with respect to patient age, general health, clinical stage and patient preference.[121] All patients had estimated survival >1 yr.
Royal Marsden Hospital, UK[92] [15839912] 1993-2002	WW	–	Any	≤7	–	–	any T stage, N0/X, M0/X	Unsuitable for RP typically because advanced age or comorbidities.
PIVOT, US [120] [18783735] 1994-2002	WW	≤75	≤50	Any	–	Bone scan negative	T1-T2/Nx/M0	–
Watchful Waiting Study, US[125] [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
Hospitals in Manchester region[xiii], UK[127] [11711356] NR	WW	–	<50	–	–	Bone scan negative	–	–
Univ. of Florida, US[123] [18263992] 2003-2006	EM	–	–	–	–	–	"Low-stage, low-grade disease"	Severe medical condition with a life expectancy of <10 yr

[xii] ERSPC-screening protocol, carried over to AS selection criteria because patients were identified from ERSPC
[xiii] University Hospital of South Manchester, Withington Hospital, Christie Hospital; Hope Hospital
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

Table 2.7. 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[114] [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[115] [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[119] [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[118] [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[116] [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.
Taichung Veterans hospital, Taiwan[117] [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[124] [12226148]	Every 6 mo in the first 2 yr, then every 1 yr	√	–	Rebiopsy was not routinely undertaken[128]	A bone scan and chest radiograph were obtained annually until	–	Hemoglobin, creatinine, alkaline phosphatase	Adjuvant local or systemic treatment was not given. TURP was as a treatment for local progression.[xiv]

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
1989-1999					1997; thereafter, chest radiographs were obtained annually for the first 2 yr			
Erasmus Univ. hospital, Netherlands[121] [7544841] ≤1990; 1993-2006[122]	Usually Followed clinically every 6 mo; Follow-up regimens varied among local practices[122]	√	–	–	"Bone scan and chest x-ray were repeated regularly"	–	Alkaline phosphatase	Local[xv] and metastatic progression were evaluated. Subjective progression, like obstructive micturition or pain, was considered for treatment decisions.[121] 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.[122]
Royal Marsden Hospital, UK[92] [15839912] 1993-2002	Every 6 mo PSA and DRE	√	–	–	–	–	–	Symptomatic prostate cancer progression
PIVOT, US [120] [18783735] 1994-2002	Every 6 mo PSA	√	–	–	Bone scan every 5 yr	–	–	Discouraged treatment for asymptomatic progression (eg, per PSA)
Watchful Waiting Study, US[125] [14501381] 1998-2003	Every 3 mo PSA	√	–	–	–	–	–	Developing progressive disease
Hospitals in Manchester, UK[127] [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[123] [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

---

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

Table 2.8. 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																					
Cohort 1[114]	PAP; PSA (after 1985)	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*
	DRE	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*	√*
	Bone scan				√				√				√				√				√
Cohort 2[125]	PAP			√				√				√				√				√	
	DRE			√				√				√				√				√	
	Bone scan			√				√				√				√				√	
Cohort 3[119]	PAP		√		√		√		√			√				√				√	
	DRE		√		√		√		√			√				√				√	
	Rebiopsy				√				√				√				√				√
	Bone scan					√				√				√							
Cohort 4[118]	PAP	monitoring frequency was not reported																			
	DRE																				
	Bone scan		√		√		√		√		√		√		√		√		√		√
Cohort 5[116]	PAP	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE																				
	Bone scan; X-ray		√		√		√		√		√		√		√		√		√		√
Cohort 6[117]	PSA (after 1990)		√		√		√		√			√				√				√	
	DRE		√		√		√		√			√				√				√	
	Bone scan																				
PSA screening era																					
Cohort 7[121]	PSA		√**		√**		√**		√**		√**		√**		√**		√**		√**		√**
	DRE																				
	Bone scan; X-ray	"repeat regularly"																			
Cohort 8[111]	PAP; PSA		√		√		√		√				√				√				√
	DRE																				
	Rebiopsy	not routine																			
	Bone scan; X-ray				√				√				√				√				√
Cohort 9[92]	PSA		√		√		√		√		√		√		√		√		√		√
	DRE		√		√		√		√		√		√		√		√		√		√
	Bone scan																				
Cohort 10[120]	PSA		√		√		√		√		√		√		√		√		√		√
	DRE																				

WW cohort	Year	1				2				3				4				5			
	Month	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60
	Bone scan																				√
Cohort 11[125]	PSA	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	DRE																				
	Bone scan																				
Cohort 12[127]	PSA		√		√		√		√		√		√		√		√		√		√
	DRE																				
	Bone scan	Probably not routine [b]																			
Cohort 13[123]	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

Cohort 1 = Howard University College of Medicine; Cohort 2 = Orebro Medical Center; Cohort 3 = Northern Stockholm region in UK; Cohort 4, Freeman Hospital, Cohort 5 = Western General Hospital; Cohort 6 = Taichung Veterans Hospital; Cohort 7 = Erasmus University Hospital, Cohort 8 = SPCG-4 trial; Cohort 9 = Royal Marsden Hospital (before 2002); Cohort 10 = PIVOT; Cohort 11 = Watchful Waiting Study; Cohort 12 = Hospitals in Manchester region in UK; Cohort 13 = University of Florida

[a] Cohorts of watchful waiting in chronological order of starting enrollment year. See Table 2.7 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 2.9). 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 2.10 and 2.11. None of these cohorts used parameters that have not been previously described.

Table 2.9. 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/Conclusions

We identified 15 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. Thirteen cohorts that enrolled only patients with clinical stages T2 or less included regular PSA testing in the followup protocol. Gleason score was also the most commonly used parameter 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 15 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 values. AS protocols generally did not include imaging in their followup protocols. In contrast, other observational strategies typically included imaging in their followups, specifically bone scan and chest radiography. They also generally did not use rebiospy but they did use PSA in their followups. Comparison of the followup frequencies between AS and other observational strategies (Tables 2.4 versus Table 2.8) 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 2.10. 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[129] [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[130] [11851612] 1991-2000	WW	–	–	–	6 sextant biopsy	–	"clinically localized prostate cancer"	–
Univ. of North Carolina, US[131] 1991-1996	EM	–	–	–	–	–	T1c	–
Princess Margaret Hospital, Canada[132] [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[133] [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[134] [21172105]] 2004-2009	AS[xvi]	–	<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

[xvi] AS criteria were created explicitly for the analyses only, not really a real AS cohort

Table 2.11. 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[129] [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[130] [11851612] 1991-2000	"a DRE", generally seen every 3-6 mo "as clinical circumstances dictated".	–	–	–	Annual bone scan	–	–	NR
Univ. of North Carolina, US[131] 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[132] [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[xvii]	√	–	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%.
BCCA,	PSA generally every 3-6 mo	√	–	–	–	–	–	NR

Center, Country [Pubmed ID] Enrollment years	Monitoring schedule	PSA	Gleason score	# biopsy cores /% cores	Imaging	Behavioral indication	Additional laboratory tests	Triggers for interventions
Canada[133] [9445192] NR	as needed							
Kansas City VA, US[134] [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

[xvii] 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.

Twenty-two 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.[67,78,135-154] These analyses were mainly conducted on the CaPSURE or SEER databases. In addition, 14 survey or interview type studies explored similar associations.[87,155-167] 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.[168] 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.[148,149] 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.

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 forego 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. 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.[157]

Acceptance of AS

No study specifically examined how a patient's primary care physician might affect the acceptance of AS. However, three survey/interview-type studies (sample sizes were 25,[156] 102,[160] and 185[87]) reported that physician recommendation (urologists or radiation oncologists) was the most influential factor in a patient's decision (30 percent in Holmboe 2000;[160] 73 percent in Gorin 2011;[87] no quantitative data were available in Davison 2009[156]) to elect or not elect AS. Two other surveys of men with prostate cancer (sample sizes were 654[158] and 231[163]) reported that physician recommendation (urologists, radiation oncologists, and others) was most influential in reaching a treatment decision (51 percent[158] and 57 percent[163]).

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.[159] 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.

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.[87]

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...."[167]

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).[165]

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.[166] 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.[155] 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.

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.[161] The responses of urologists and radiation oncologists did not differ significantly.

Acceptance of AS

Age. Ten multivariable analyses provided results for age with respect to AS/WW.[67,78,135,137,138,142,147,150,151,153] 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.[67,78,135,140-142] However, one analysis found no such association.[153]

Gleason score. Two multivariable analyses reported that men with a higher Gleason score also had a decreased probability of receiving AS/WW versus any other treatment.[78] or RP.[147]

Histopathology. Two multivariable analyses provided results for histopathology with respect to AS/WW.[135,142] 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;[169] T2 versus T1[147,153]) 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.[78,147,153]

Risk groups. Two multivariable analyses reported that men assessed as D'Amico low-risk[170] (versus intermediate- or high-risk) had an increased probability of receiving AS/WW versus active treatments.[67,137]

Adherence to AS

Age. Four multivariable analyses analyzed age with respect to interruption of AS/WW.[140,144,151,152] Two found that men who were younger had an increased probability of receiving definitive treatments,[144,151] and two found that age was not a factor.[140,152]

Comorbidities. Three multivariable analyses reported that the number of comorbidities was not associated with interruption of AS/WW.[140,151,152]

Gleason score. Five multivariable analyses reported that Gleason score at diagnosis was not associated with interruption of AS/WW.[144,148,150-152]

Histopathology. No study analyzed histopathology 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.[144,151,152] One analysis reported that disease stage (T2 versus T1) was not a significant predictor of the receipt of curative treatment.[150]

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.[144,151] One analysis reported that neither initial PSA nor PSA density was predictive of men who interrupted AS to seek radical treatment, but an increased free to total PSA ratio was predictive of the probability of interruption of AS.[148] 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).[152] Two other analyses also reported that increased PSA velocity in men was predictive of interruption of AS to seek active treatments.[140,150]

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,[143] while two reported that risk classification was not a significant predictor in the interruption of AS/WW.[140,144]

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,[158] and another survey 9 percent of 231 men,[163] 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.[164] 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).[160]

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 selected the AS arm with 138 men who were randomized.[149] 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).

Three survey/interview type studies (sample size 25,[156] 50,[162] and 185[87]) reported that concern for treatment side effects (impotence [44 percent] and incontinence [48 percent];[87] no quantitative data were available in 2 studies[156,162]) was one reason that patients elected AS/WW.

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.[154] 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.[153] Another multivariable analysis reported that having other urinary conditions (besides the primary urinary dysfunction) was also predictive of choosing WW over RP.[141]

Adherence to AS

One multivariable analysis reported that increased anxiety in men was associated with an increased probability of interruption of AS.[140]

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.[156] 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.[158] Another survey of 102 men with localized prostate cancer reported that fear of consequences was the most common reason for not selecting WW.[160]

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.[151,152]

Social support

Offer of AS

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

Acceptance of AS

Three multivariable analyses reported that not being married or in a permanent relationship was associated with an increased probability of receiving WW versus active treatments.[78,136,142] One analysis found that marital status was not a factor in predicting receiving WW.[153]

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.[167] 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.[140]

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.[67,153,169] 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.[169] The remaining two reported that income was not a significant factor in predicting the choice of AS/WW versus other treatments.[67,153]

One multivariable analysis reported that men in higher socioeconomic strata (versus lower strata) had a decreased probability of receiving AS/WW versus active treatments.[78,135]

Education. Five multivariable analyses examined education level with respect to AS/WW.[67,137,142,143,153] Three reported that education was not a significant factor in predicting men receiving AS/WW.[67,137,153] 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.[142] One reported that college graduates (versus non-college graduates) had an increased probability of interrupting AS/WW to seek active treatments.[143]

Race. Nine multivariable analyses examined race/ethnicity with respect to AS/WW.[78,135,140,142,144-146,152,153] 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.[78,140,144-146,152] Three analyses reported that blacks were more likely than whites to receive AS/WW versus active treatments.[135,142,153]

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).[163]

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,[67] or private or Veterans Administration (VA) insurance[136]) increased the probability of receiving AS/WW. 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.[139] 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.[140]

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%).[163]

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 U.S. (ranging from 0 to 28 percent) and that this variation was not explained by known patient factors.[70] 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.[78] 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).

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.[135] 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). [163]

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 vs. 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.[67]

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[171]) 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.[168] 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,[172] the reviewers rated two RCTs as good[173,174] and one poor.[175]

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.[175] 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.[174] One RCT found that the men in the DA arm selected their physician's treatment choice less often than those who received usual care.[173] 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.[176]

Summary/Conclusions

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).[148,149] 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.[148] 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).[149]

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, higher Gleason score, higher tumor stage, higher diagnostic PSA, membership in a higher 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?

Clinical Outcomes

We did not identify any studies reporting clinical outcomes specifically of AS management strategies including deferred treatment with curative intent versus 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. For completeness, though we have summarized the evidence from randomized and nonrandomized comparative studies that used other observational management strategies that used no immediate active treatment and mostly resembled WW.

Findings from previous systematic reviews

We examined two recent systematic reviews of treatments of men with clinically localized prostate cancer.[5,6] Of note, these systematic reviews included studies of men receiving ADT in their observational management groups, studies that would have been rejected in the systematic review for the current report. One review[5] that compared RT and no treatment evaluated one prospective cohort study[177] and eight[178-185] 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.[177] 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.[179] 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.[178] One study reported significantly higher rates of second primary cancer in patients treated with EBRT compared with those with no treatment.[180]

Another review[6] included two RCTs comparing WW with RP: the Scandinavian Prostate Cancer Group Study 4 (SPCG-4),[186] and the Veterans Administration Cooperative Urological Research Group (VACURG) trial.[187] A recent (2010) Cochrane Report on the same topic[188] did not identify any additional studies; however, we identified the latest update of the SPCG-4 trial results[189] (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).[186] The VACURG trial followed 142 patients for a median of 23 years, and found no difference in mortality between WW and RP groups.[187]

Additionally, the AHRQ evidence report on localized prostate cancer treatment[6] considered the results of two randomized trials reporting on QoL and self-reported functional status.[190,191] One included study was an ancillary investigation from the SPCG-4[190]. 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.[190] The second study was based on a randomized trial comparing RT with deferred treatment[191] 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.[191]

Findings from primary studies

To address Key Question 4, we searched for studies that compared observational management strategies including deferred treatment with curative intent with immediate definitive treatment. We included only multicenter studies that enrolled men with localized prostate cancer, and reported age-adjusted effect sizes. Characteristics of the 11 eligible studies reporting on clinical outcomes are shown in Table 4.1. One RCT,[189] two prospective cohort studies,[192 193]and eight retrospective cohort studies[178,184,185,194-198] were included. The RCT is an update of the SPCG-4 trial included in the previously mentioned systematic review.[6] Among the ten cohort studies, sample size ranged from 113[193] to 44,630[197]and followup duration from 12 months[193] to more than 12 years[189] Methodological quality of the studies were rated as B in eight studies,[178,184,185,189,194-197] and C in three studies.[192,193,199]

Comparison between observational management strategies and radical prostatectomy[a]

Results from one RCT and eight cohort studies that compared observational management strategies with RP are shown in Table 4.2.

The RCT followed 695 men with localized prostate cancer for a median of 12.8 years.[189] 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".[189] 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.

---

[a] 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.

Two cohort studies[184,196] compared prostate cancer-specific mortality between patients on conservative management and patients treated with RP based on large databases (SEER-Medicare[196] and the National Cancer Register of Sweden Follow-up Study, NCRSFS[184]). Both studies identified a statistically significant difference between treatments favoring RP (HR 0.63; 95 percent CI 0.55, 0.71 [using a propensity score[196]], and HR 0.49; 95 percent CI 0.34, 0.71,[184] 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).[196]

Four studies compared all-cause mortality between observational management strategies and RP.[184,195-197] One was based on the CDC-NPCR Patterns of Care Study (POCS),[195] two used the SEER-Medicare database,[196,197] and one used NCRSFS data.[184] 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).[195] Similarly, the two reports[196,197] that analyzed data from the SEER-Medicare database and the NCRSFS study[184] showed favorable outcome in the RP group (HR 0.65; 95 percent CI 0.62 ,0.68[using a propensity score[196]]; HR 0.49; 95 percent CI 0.41, 0.57[196]; and HR 0.50; 95 percent CI 0.47, 0.53,[184] respectively).[184,196,197] 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.[196]

One report analyzed the data from the CaPSURE registry 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).[178] One study analyzed the risk of additional surgical procedures following primary treatment with WW, RP, RT or ADT.[198] 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 more common in the WW group (HR 0.71; 95 percent CI, 0.61-0.84) and cystoscopy procedures were equally common (HR 1.00; 95 percent CI, 1.33-2.20), compared to the RP group.

Three studies reported QoL outcomes.[192-194] From the CaPSURE registry, patients on observational management strategies (WW or AS, reported in aggregate) had significantly lower mean SF-36 score in the social function domain than patients treated with RP (89 vs. 100; $P < 0.05$, respectively), but no difference was found in three other domains of SF-36.[194] From the SEER and PCOS databases, proportions of patients who rated "satisfied" with their treatments for prostate cancer in the RP group and rated "satisfied" in the no treatment group were 57.8 percent (95 percent CI 54.1-61.5; P not available) and 50.5 percent (95 percent CI 42.5-58.8; P not reported), respectively.[192] One study from four academic centers in Wisconsin 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.[193]

Comparison between observational management strategies and radiation therapy

Results from eight studies that compared observational management strategies and radiation therapy (RT) are shown in Table 4.3. No study reported on the effects of treatment on the development of metastatic disease.

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.

Two studies compared reported adjusted estimates comparing all-cause mortality between observational management strategies and RT.[184,197] One study was based on the SEER-Medicare database,[197] , and the other used NCRSFS data.[184] 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,[197] and HR 0.68; 95 percent CI 0.57, 0.82[184], respectively).

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.[178]

Four studies reported quality of life outcomes.[185,192-194] 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.[194] From the SEER and PCOS databases, proportions of patients who rated "satisfied" with their treatments for prostate cancer in the RT group and rated "satisfied" in the no treatment group were 69.4 percent (95 percent CI 64.6-74.2) and 50.5 percent (95 percent CI 42.5-58.8), respectively. No P value was reported for this comparison.[192] 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.[193] 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

Results from 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 Table 4.4. 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.[197] 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).[197]

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).[178] No significant difference was found between patients on WW and patients treated with combined RP and EBRT.[178]

Costs

We identified three primary studies (two using U.S. data[135,200] and one Sweden[201]) 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 C4.1.

One study used the SEER-Medicare database (13,769 patients matched 1:1 with controls; 2805 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.[135] A second study used data from the CaPSURE database (235 patients; 37 managed with WW) to estimate mean first year costs.[200] 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 was an ancillary investigation from the Scandinavian Prostatic Cancer Group Study Number 4, a randomized trial of RP versus WW for men with localized prostate cancer.[201] 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 ($P = 0.003$).

In addition to the primary cost studies discussed above, we also considered two economic evaluations conducted by the Institute for Clinical and Economic Reviews (ICER)[7,8] along with an associated publication based on these reports.[9] We did not consider the cost-effectiveness analyses reported in these documents and focused only on cost information. In both ICER reports, 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. In the base case analysis (cohort of 65 year old men diagnosed with low-risk clinically localized prostate cancer, followed for 15 years, 3 percent annual discount rate) comparing AS (followed by RT in cases of progression or patient preference) 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, followed for 15 years, 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.

Summary/Conclusions

No study reported clinical outcomes specifically for AS management strategies including deferred treatment with curative intent 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 and 10 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 previously mentioned RCT 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 and satisfaction with treatment 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 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 suggests that AS may be associated with higher costs compared to RP or BT, but lower costs compared to intensity modulated RT (IMRT) or proton beam RT.

Table 4.1. 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 study – clinical outcomes							
Bill-Axelson, 2011[189] [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 well differentiated tumor	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%	B
Randomized study – treatment costs							
Andersson, 2011[201] [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
Observational studies – clinical outcomes							
Stattin, 2010[184]	NPCRSFS	Surveillance (AS	Median	6849	≤70 yr, clinically localized disease (T1/2),	Mean age: surveillance, 64.7 yr; RP, 61.2;	B

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
[20562373] Retrospective cohort		and WW, in aggregate) vs. RP vs. RT	followup = 8.2 yr		N0/x, M0/x, PSA<20 ng/mL; Gleason score ≤7	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%.	
Litwin, 2002[194] [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
Schymura, 2010[195] [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
Hadley, 2010[196] [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	14302 used in survival models [a sample of 17815 patients was used for PS and IV analyses; then, exclusion of	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	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

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
			Dec 31, 2007 (censoring date) was 78 mo (IQR=48 mo)	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]	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.		
Wong, 2006[197] [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-diff., 25.87%; moderately diff., 64.13%. Active treatment, well-diff., 14.29%; moderately dif., 85.71% Observation: ≤T2a, 55.03%; T2b/c, 44.97%; active treatment, ≤T2a, 37.92%; T2b/c, 62.08%	B
Berge, 2007[198] [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	Median age: WW, 77; RP, 70; RT, 74 PSA: NR Grade: WW, well dif., 40.2%; moderately	C

[a] We did not extract data from the group of patients receiving primary ADT as the only initial therapy.

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
					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.	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%.	
Elliott, 2007[178] [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
Hoffman, 2003[192] [12655522] Prospective cohort	PCOS/SEER	NT vs. RT vs. RP	2 yr	2365	completed 2 yr f/u ; T1/2 tumors	Age: 66 yr (39-88) PSA: <4 : 11% ; 4-9.9 : 57% ; ≥10 : 32% Gleason score: 2-4 : 19% ; 5-7 : 74% ; 8-10 : 7% Stage: NR	C Survey responders and nonresponders were different in baseline characteristics.
Schapira, 2001[193] [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	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-	C Selection bias: 19% eligible patients were not contacted for a variety of reasons; dropout rate

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
					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)	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%	12%, 9%, 7% in RP, RT, and RM group, respectively
Thong, 2009[185] [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%)
Observational studies – treatment costs							
Snyder, 2010[135] [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%	C

[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.

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
						Stage: NR (clinically localized 100%)	
Penson, 2001[200] [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

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.

Table 4.2. Comparison between active surveillance or 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, 2011[189] [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, 2010[196] [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.
Stattin, 2010[184] [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.
All-cause							

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
mortality							
Bill-Axelson, 2011[189] [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, 2010[195] [20403178] Retrospective cohort	POCS CDC-NPCR	WW. vs. RP	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
Hadley, 2010[196] [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, 2010[184] [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.
Wong, 2006[197] [17164454] Retrospective cohort	SEER-Medicare	RP vs. observation	Overall survival = interval from the date of diagnosis to the Medicare dare	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

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
			of death. Patients were censored at Dec. 20, 2002.				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.
Incidence of distant metastases							
Bill-Axelson, 2011[189] [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							
Elliott, 2007[178] [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
Berge, 2007[198] [17178188]	SEER-Medicare	WW vs. RP[c]	Cystoscopy; bladder	5 yr	WW: 3612 RP: 3940	HR 1.00 (0.88, 1.13) for cystoscopy; P=0.071 for the null hypothesis that the	Age, grade, comorbidity index

[c] 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.

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
Retrospective cohort			irrigation/ cystostomy; TURP/bladder-neck incision; urethra dilation [procedures considered indicative of treatment-related morbidity]			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	
Quality of life							
Litwin, 2002[194] [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
Hoffman, 2003[192] [12655522] Prospective cohort	PCOS/SEER	NT vs. RP vs. RT vs. ADT	percentage distributions satisfied with treatment	2 yr	NT: 230 RT: 583 RP: 1373 ADT: 179	NT 50.5% (42.5, 58.8) RT 69.4% (64.6, 74.2) RP 57.8% (54.1, 61.5) ADT 66.3% (58.0, 74.6) Wald chi-square test P <0.001	Age, race, geographic area, current disease status, previous symptoms, treatment morbidity, general health, and impact of cancer on activity and relationship with spouse/friends; weighted to the total population.
Schapira, 2001[193] [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	Patient age, comorbidity, TNM, PSA, race, marital status, working status, and years of education.

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
						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.	

Dec = December; EM = expectant management; HR = hazard ratio; IPTW = inverse probability of treatment weights; IQR = interquartile range; IV = instrumental variable; mo = months; 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; SMRW = standardized mortality ratio weights; TURP = transurethral resection of the prostate; WW = watchful waiting; yr = year

Table 4.3. Comparison between active surveillance or 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, 2010[184] [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, 2006[197] [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.
Stattin, 2010[184] [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.
Morbidity of							

primary treatment							
Elliott, 2007[178] [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
Quality of life							
Litwin, 2002[194] [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
Hoffman, 2003[192] [12655522] Prospective cohort	PCOS/SEER	NT vs. RP vs. RT vs. ADT	percentage distributions satisfied with treatment	2 yr	NT: 230 RT: 583 RP: 1373 ADT: 179	NT 50.5% (42.5, 58.8) RT 69.4% (64.6, 74.2) RP 57.8% (54.1, 61.5) ADT 66.3% (58.0, 74.6) Wald chi^2 test P <0.0001	Age, race, geographic area
Schapira, 2001[193] [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.
Thong, 2009[185] [19747357] Retrospective matched cohort	Eindhoven Cancer Registry (ECR)	RT vs. "AS"	General QoL: SF-36 scores Expanded Prostate Cancer Index (EPIC): urinary	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.	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

	and bowel functioning, and urinary and bowel bother Quality of Life – Cancer Survivors (QOL-CS)	No other significant associations between general QoL, cancer-specific QoL, or disease-specific QoL scores and management strategy (RT vs. "AS").	progression

ADT = androgen deprivation therapy; BT = brachytherapy; HR = hazard ratio; NR = not reported; RP = radical prostatectomy; RT = radiation therapy.

Table 4.4. Comparison between active surveillance or 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, 2006[197] [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, 2006[197] [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.
Morbidity of primary treatment							
Elliott, 2007[178] [17570425] Retrospective cohort	CaPSURE	RP+EBRT (combination treatment) vs. WW	Treatment for urethral stricture identified by study abstracted hospital records	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

		includes (ICD codes)*			
	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

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

As summarized and discussed above, 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 active surveillance or of watchful waiting, the standard protocols for the interventions, or how to manage patients whose cancers show signs of progression. Thus, it is difficult when reviewing studies to know which patients had true active surveillance or watchful waiting, or who were simply not treated (for a variety of reasons), or who had delays in their treatment (and thus initially had no treatment). Furthermore, it has been 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.

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 stage and grade information are often incomplete 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 interventions. 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 report summary stage information using the "a combination of the most precise clinical and pathological documentation of the extent of disease".[a] This may result in

[a] See the SEER staging manual, available at: http://seer.cancer.gov/tools/ssm/ (last accessed: August 7, 2011).

misclassification, as the accuracy of available information on staging depends on the treatment patients receive. Those who had RP will have pathological staging information, in contrast with those who receive RT, ADT alone, AS, or WW, who have clinical staging only. 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 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 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, symptomatic disease, urological and other complications (from testing or 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,[144] 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).

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, acceptance, 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 future database analyses 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.

When better data become available regarding the factors that affect the offer, acceptance, and adherence of AS as delineated in the Key Question, 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 causal 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.[202,203]

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

The least biased, most reliable study design comparing two interventions is the well-conducted randomized controlled trial that adheres to modern 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 tests, adverse events, resource utilization, and costs. However, we acknowledge that conducting and completing an adequately powered trial of sufficient duration may be challenging. 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. Trials would then need to be of sufficiently long duration to collect data on the clinically relevant outcomes.

In lieu of randomized trials, adequate findings may be possible 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*. 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.

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. 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 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 5-7). It has been hypothesized that this effect is caused by changes in histopathological grading guidelines,[204] a preference towards avoiding assigning Gleason 2-4 scores based on prostate cancer biopsy samples[85,205,206], 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 15 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, it was difficult when reviewing the studies to know which patients had true AS or WW, or who were simply not treated (for a variety of reasons), or who had 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 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, higher Gleason score, higher tumor stage, higher diagnostic PSA, higher 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 provided results from 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. 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]). One other trial[c] was stopped early because of limited enrollment. A brief description of these studies is provided in Appendix Table B.

Although costs 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.

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 active surveillance that clearly distinguishes it from watchful waiting and other observational management strategies is needed to help clarify scientific discourse in this field. Ongoing clinical trials may provide information on the

---

[a] See http://clinicaltrials.gov/ct2/show/study/NCT00007644; last accessed July 22, 2011.
[b] See http://clinicaltrials.gov/ct2/show/NCT00632983; last accessed July 22, 2011.
[c] See http://clinicaltrials.gov/ct2/show/NCT00499174; last accessed July 22, 2011.

comparative effectiveness of active surveillance compared to immediate active treatment, but will require long term followup.

References

(1) Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin.* 2011;61:212-236.

(2) Cooperberg MR, Lubeck DP, Mehta SS, Carroll PR. Time trends in clinical risk stratification for prostate cancer: implications for outcomes (data from CaPSURE). *J Urol.* 2003;170:S21-S25.

(3) Agency for Healthcare Research and Quality. Methods Reference Guide for Effectiveness and Comparative Effectiveness Reviews, Version 1.0. 2010.
Ref Type: Generic

(4) Martin RM, Gunnell D, Hamdy F, Neal D, Lane A, Donovan J. 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.

(5) Ip, S., Dvorak, T., Yu, W. W., Patel, K., Obadan, N. O., Chung, M., Bannuru, R. R., and Lau, J. Comparative evaluation of radiation treatments for clinically localized prostate cancer: an update http://www.cms.gov/mcd/viewtechassess.asp?where=index&tid=69. 8-13-2010.
Ref Type: Report

(6) Wilt TJ, MacDonald R, Rutks I, Shamliyan TA, Taylor BC, Kane RL. Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. *Ann Intern Med.* 2008;148:435-448.

(7) Ollendorf, D. A., Hayes, J. H., McMahon, P. M., and Pearson, S. D. Active surveillance & radical prostatectomy for the management of low-risk, clinically-localized prostate cancer. 1-11-2009. Institute for Clinical and Economic Review.
Ref Type: Report

(8) Ollendorf, D. A., Hayes, J. H., McMahon, P. M., Kuba, M., and Pearson, S. D. Management options for low-risk prostate cancer. 2009. Insitute for Clinical and Economic Review.
Ref Type: Report

(9) 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.

(10) Moher D, Liberati A, Tetzlaff J, Altman DG. 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, Cummings KB, Ward WS. 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, Ward WS, Cummings KB. 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, Fay MP, Hankey BF. 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, Tonita J, Thompson TD. 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, Edwards BK. 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, Master VA, Brawley OW. 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, Wali P, Kramer B. 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, Liauw SL, Johnstone PA. 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, Miller BA, 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, Montie JE, Wei JT. 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, Friedman C, Wingo PA. 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 W, III, 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, Fagerlin A, Sanda MG, Wei JT. Racial/ethnic disparities in the treatment of localized/regional prostate cancer. *J Urol.* 2004;171:1504-1507.
(49) Demers RY, Swanson GM, Weiss LK, Kau TY. 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, Demers RY. 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, Weiss LK, Severson RK, Montie J. 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, Wood DP, Jr. 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, Stone N, Harlan LC, Key CR. 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, Brawer MK. 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, James BC, Janerich DT, Dibble RL. 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, Kramer BS. The role of increasing detection in the rising incidence of prostate cancer. *JAMA.* 1995;273:548-552.
(58) Klabunde CN, Potosky AL, Harlan LC, Kramer BS. 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, Ramsey SD, Penson DF. 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, Ross LE, Wobker SE, Godley PA. Racial differences in PSA screening interval and stage at diagnosis. *Cancer Causes Control.* 2010;21:1071-1080.
(63) Mullins CD, Onukwugha E, Bikov K, Seal B, Hussain A. 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, Flanders SC, Henning JM, Carroll PR. 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, Mehta SS, Carroll PR. 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) Cooperberg MR, Grossfeld GD, Lubeck DP, Carroll PR. National practice patterns and time trends in androgen ablation for localized prostate cancer. *J Natl Cancer Inst.* 2003;95:981-989.
(67) 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.

(68) Cooperberg MR, Lubeck DP, Meng MV, Mehta SS, Carroll PR. The changing face of low-risk prostate cancer: trends in clinical presentation and primary management. *J Clin Oncol.* 2004;22:2141-2149.
(69) Cooperberg MR, Broering JM, Kantoff PW, Carroll PR. Contemporary trends in low risk prostate cancer: risk assessment and treatment. *J Urol.* 2007;178:S14-S19.
(70) 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.
(71) Shah JB, McKiernan JM, Elkin EP, Carroll PR, Meng MV. Prostate biopsy patterns in the CaPSURE database: evolution with time and impact on outcome after prostatectomy. *J Urol.* 2008;179:136-140.
(72) Greene KL, Cowan JE, Cooperberg MR, Meng MV, DuChane J, Carroll PR. Who is the average patient presenting with prostate cancer? *Urology.* 2005;66:76-82.
(73) Mettlin CJ, Murphy G. The National Cancer Data Base report on prostate cancer. *Cancer.* 1994;74:1640-1648.
(74) Mettlin CJ, Murphy GP, Ho R, Menck HR. The National Cancer Data Base report on longitudinal observations on prostate cancer. *Cancer.* 1996;77:2162-2166.
(75) Mettlin CJ, Murphy GP, McGinnis LS, Menck HR. 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.
(76) Mettlin CJ, Murphy GP, Rosenthal DS, Menck HR. 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.
(77) Danley KL, Richardson JL, Bernstein L, Langholz B, Ross RK. 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.
(78) 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.
(79) Frey CM, McMillen MM, Cowan CD, Horm JW, Kessler LG. Representativeness of the surveillance, epidemiology, and end results program data: recent trends in cancer mortality rates. *J Natl Cancer Inst.* 1992;84:872-877.
(80) 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.
(81) Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care.* 2002;40:IV-18.
(82) 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.
(83) Albertsen PC, Hanley JA, Barrows GH et al. Prostate cancer and the Will Rogers phenomenon. *J Natl Cancer Inst.* 2005;97:1248-1253.
(84) 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.

(85) Gofrit ON, Zorn KC, Steinberg GD, Zagaja GP, Shalhav AL. The Will Rogers phenomenon in urological oncology. *J Urol.* 2008;179:28-33.
(86) 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.
(87) Gorin MA, Soloway CT, Eldefrawy A, Soloway MS. Factors that influence patient enrollment in active surveillance for low-risk prostate cancer. *Urology.* 2011;77:588-591.
(88) 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.
(89) Patel MI, DeConcini DT, Lopez-Corona E, Ohori M, Wheeler T, Scardino PT. An analysis of men with clinically localized prostate cancer who deferred definitive therapy. *J Urol.* 2004;171:1520-1524.
(90) Soloway MS, Soloway CT, Williams S, Ayyathurai R, Kava B, Manoharan M. Active surveillance; a reasonable management alternative for patients with prostate cancer: the Miami experience. *BJU Int.* 2008;101:165-169.
(91) Soloway MS, Soloway CT, Eldefrawy A, Acosta K, Kava B, Manoharan M. 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.
(92) Hardie C, Parker C, Norman A et al. Early outcomes of active surveillance for localized prostate cancer. *BJU Int.* 2005;95:956-960.
(93) Venkitaraman R, Norman A, Woode-Amissah R et al. Prostate-specific antigen velocity in untreated, localized prostate cancer. *BJU Int.* 2008;101:161-164.
(94) 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.
(95) 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.
(96) Klotz L, Zhang L, Lam A, Nam R, Mamedov A, Loblaw 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.
(97) 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.
(98) San F, I, Werner L, Regan MM, Garnick MB, Bubley G, DeWolf WC. 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.
(99) Miocinovic R, Jones JS, Pujara AC, Klein EA, Stephenson AJ. 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.
(100) 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.

(101) 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.
(102) Fryback DG, Albertsen PC, Storer BE. Prostatectomy and survival among men with clinically localized prostate cancer. *JAMA.* 1996;276:1723-1724.
(103) 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.
(104) Ercole B, Marietti SR, Fine J, Albertsen PC. Outcomes following active surveillance of men with localized prostate cancer diagnosed in the prostate specific antigen era. *J Urol.* 2008;180:1336-1339.
(105) 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.
(106) 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.
(107) 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.
(108) Neulander EZ, Duncan RC, Tiguert R, Posey JT, Soloway MS. 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.
(109) 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.
(110) 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.
(111) 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.
(112) 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.
(113) 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.
(114) Jones GW. Prospective, conservative management of localized prostate cancer. *Cancer.* 1992;70:307-310.
(115) Johansson JE. Expectant management of early stage prostatic cancer: Swedish experience. *J Urol.* 1994;152:1753-1756.
(116) Rana A, Chisholm GD, Christodoulou S, McIntyre MA, Elton RA. Audit and its impact in the management of early prostatic cancer. *Br J Urol.* 1993;71:721-727.
(117) 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.

(118) Handley R, Carr TW, Travis D, Powell PH, Hall RR. Deferred treatment for prostate cancer. *Br J Urol.* 1988;62:249-253.
(119) 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.
(120) 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.
(121) 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.
(122) 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.
(123) Anai S, Nakamura K, Chang MN, Pendleton J, Yacoub S, Rosser CJ. The feasibility of expectant management with inner-city men with newly diagnosed localized prostate cancer. *J Health Care Poor Underserved.* 2008;19:164-170.
(124) 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.
(125) 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.
(126) 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.
(127) McIntyre IG, Clarke RB, Anderson E, Clarke NW, George NJ. Molecular prediction of progression in patients with conservatively managed prostate cancer. *Urology.* 2001;58:762-766.
(128) 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.
(129) 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.
(130) Arai Y, Egawa S, Kuwao S, Ogura K, Baba S. 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.
(131) 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.
(132) 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.

(133) McLaren DB, McKenzie M, Duncan G, Pickles T. 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.
(134) Lee EK, Baack J, Penn H et al. Active surveillance for prostate cancer in a veteran population. *Can J Urol.* 2010;17:5429-5435.
(135) 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.
(136) Dall'era MA, Hosang N, Konety B, Cowan JE, Carroll PR. Sociodemographic predictors of prostate cancer risk category at diagnosis: unique patterns of significant and insignificant disease. *J Urol.* 2009;181:1622-1627.
(137) Barocas DA, Cowan JE, Smith JA, Jr., Carroll PR. 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.
(138) 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.
(139) Sadetsky N, Elkin EP, Latini DM, DuChane J, Carroll PR. Prostate cancer outcomes among older men: insurance status comparisons results from CaPSURE database. *Prostate Cancer Prostatic Dis.* 2008;11:280-287.
(140) 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.
(141) Marr PL, Elkin EP, Arredondo SA, Broering JM, DuChane J, Carroll PR. Comorbidity and primary treatment for localized prostate cancer: data from CaPSURE. *J Urol.* 2006;175:1326-1331.
(142) 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.
(143) Meng MV, Elkin EP, Harlan SR, Mehta SS, Lubeck DP, Carroll PR. Predictors of treatment after initial surveillance in men with prostate cancer: results from CaPSURE. *J Urol.* 2003;170:2279-2283.
(144) 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.
(145) Moses KA, Paciorek AT, Penson DF, Carroll PR, Master VA. Impact of ethnicity on primary treatment choice and mortality in men with prostate cancer: data from CaPSURE. *J Clin Oncol.* 2010;28:1069-1074.
(146) Latini DM, Elkin EP, Cooperberg MR, Sadetsky N, DuChane J, Carroll PR. 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.
(147) 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.

(148) 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.
(149) 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.
(150) 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.
(151) 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.
(152) 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.
(153) Yan Y, Carvalhal GF, Catalona WJ, Young JD. Primary treatment choices for men with clinically localized prostate carcinoma detected by screening. *Cancer.* 2000;88:1122-1130.
(154) 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.
(155) 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.
(156) Davison BJ, Oliffe JL, Pickles T, Mroz L. Factors influencing men undertaking active surveillance for the management of low-risk prostate cancer. *Oncol Nurs Forum.* 2009;36:89-96.
(157) Durham J, Low M, McLeod D. Screening for prostate cancer: a survey of New Zealand general practitioners. *N Z Med J.* 2003;116:U476.
(158) Diefenbach MA, Dorsey J, Uzzo RG et al. Decision-making strategies for patients with localized prostate cancer. *Semin Urol Oncol.* 2002;20:55-62.
(159) Zietman AL, Thakral H, Wilson L, Schellhammer P. 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.
(160) Holmboe ES, Concato J. Treatment decisions for localized prostate cancer: asking men what's important. *J Gen Intern Med.* 2000;15:694-701.
(161) Fowler FJ, Jr., McNaughton CM, Albertsen PC, Zietman A, Elliott DB, Barry MJ. Comparison of recommendations by urologists and radiation oncologists for treatment of clinically localized prostate cancer. *JAMA.* 2000;283:3217-3222.
(162) Chapple A, Ziebland S, Herxheimer A, McPherson A, Shepperd S, Miller R. Is 'watchful waiting' a real choice for men with prostate cancer? A qualitative study. *BJU Int.* 2002;90:257-264.
(163) 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.

(164) Berry DL, Ellis WJ, Woods NF, Schwien C, Mullen KH, Yang C. Treatment decision-making by men with localized prostate cancer: the influence of personal factors. *Urol Oncol.* 2003;21:93-100.
(165) Steginga SK, Occhipinti S, Gardiner RA, Yaxley J, Heathcote P. Making decisions about treatment for localized prostate cancer. *BJU Int.* 2002;89:255-260.
(166) Crawford ED, Bennett CL, Stone NN et al. Comparison of perspectives on prostate cancer: analyses of survey data. *Urology.* 1997;50:366-372.
(167) O'Rourke ME. Narrowing the options: the process of deciding on prostate cancer treatment. *Cancer Invest.* 1999;17:349-359.
(168) Lin GA, Aaronson DS, Knight SJ, Carroll PR, Dudley RA. Patient decision aids for prostate cancer treatment: a systematic review of the literature. *CA Cancer J Clin.* 2009;59:379-390.
(169) 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.
(170) 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.
(171) Shariat SF, Karakiewicz PI, Roehrborn CG, Kattan MW. An updated catalog of prostate cancer predictive tools. *Cancer.* 2008;113:3075-3099.
(172) 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.
(173) 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.
(174) Davison BJ, Goldenberg SL, Wiens KP, Gleave ME. Comparing a generic and individualized information decision support intervention for men newly diagnosed with localized prostate cancer. *Cancer Nurs.* 2007;30:E7-15.
(175) Davison BJ, Degner LF. Empowerment of men newly diagnosed with prostate cancer. *Cancer Nurs.* 1997;20:187-196.
(176) 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.
(177) Choo R, Long J, Gray R, Morton G, Gardner S, Danjoux C. 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.
(178) Elliott SP, Meng MV, Elkin EP, McAninch JW, DuChane J, Carroll PR. Incidence of urethral stricture after primary treatment for prostate cancer: data From CaPSURE. *J Urol.* 2007;178:529-534.
(179) 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.
(180) 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.

(181) Albertsen PC, Hanley JA, Penson DF, Barrows G, Fine J. 13-year outcomes following treatment for clinically localized prostate cancer in a population based cohort. *J Urol.* 2007;177:932-936.
(182) 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.
(183) 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.
(184) Stattin P, Holmberg E, Johansson JE, Holmberg L, Adolfsson J, Hugosson J. Outcomes in localized prostate cancer: National Prostate Cancer Register of Sweden follow-up study. *J Natl Cancer Inst.* 2010;102:950-958.
(185) Thong MS, Mols F, Kil PJ, Korfage IJ, van de Poll-Franse LV. 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.
(186) 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.
(187) 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.
(188) Hegarty J, Beirne PV, Walsh E, Comber H, Fitzgerald T, Wallace KM. Radical prostatectomy versus watchful waiting for prostate cancer. *Cochrane Database Syst Rev.* 2010;CD006590.
(189) 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.
(190) 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.
(191) Fransson P, Damber JE, Tomic R, Modig H, Nyberg G, Widmark A. Quality of life and symptoms in a randomized trial of radiotherapy versus deferred treatment of localized prostate carcinoma. *Cancer.* 2001;92:3111-3119.
(192) Hoffman RM, Hunt WC, Gilliland FD, Stephenson RA, Potosky AL. Patient satisfaction with treatment decisions for clinically localized prostate carcinoma. Results from the Prostate Cancer Outcomes Study. *Cancer.* 2003;97:1653-1662.
(193) Schapira MM, Lawrence WF, Katz DA, McAuliffe TL, Nattinger AB. Effect of treatment on quality of life among men with clinically localized prostate cancer. *Med Care.* 2001;39:243-253.
(194) Litwin MS, Lubeck DP, Spitalny GM, Henning JM, Carroll PR. 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.
(195) 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.
(196) Hadley J, Yabroff KR, Barrett MJ, Penson DF, Saigal CS, Potosky AL. Comparative effectiveness of prostate cancer treatments: evaluating statistical

adjustments for confounding in observational data. *J Natl Cancer Inst.* 2010;102:1780-1793.
(197) 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.
(198) 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.
(199) Berge V, Thompson T, Blackman D. Use of additional treatment for prostate cancer after radical prostatectomy, radiation therapy, androgen deprivation, or watchful waiting. *Scand J Urol Nephrol.* 2007;41:198-203.
(200) 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.
(201) 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.
(202) Robins J, Orellana L, Rotnitzky A. Estimation and extrapolation of optimal treatment and testing strategies. *Stat Med.* 2008;27:4678-4721.
(203) Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol.* 2006;98:237-242.
(204) Srigley JR, Amin MB, Bostwick DG, Grignon DJ, Hammond ME. 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.
(205) 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.
(206) 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.
(207) Tseng KS, Landis P, Epstein JI, Trock BJ, Carter HB. Risk stratification of men choosing surveillance for low risk prostate cancer. *J Urol.* 2010;183:1779-1785.