

Evidence-based Practice Center Review Protocol

NIH State-of-the-Science Conference: Role of Active Surveillance in the Management of Men With Localized Prostate Cancer

I. Background and Objectives for the Review

The National Cancer Institute and the Centers for Disease Control and Prevention are sponsoring a National Institutes of Health (NIH) State-of-the-Science Conference to examine the role of active surveillance—as opposed to immediate curative intent therapy (e.g., radical prostatectomy or radiation therapy)—in the management of early stage, low-risk prostate cancer. It is anticipated that the Conference will take place on December 5–7, 2011.

Currently, most prostate cancer in men in the United States is found by prostate-specific antigen (PSA) screening. The cancer is often localized, and most tumors have low histological grades and Gleason Scores. Such cancer is an infrequent cause of death; affected patients are more likely to die of other unrelated causes. It was projected that in 2010 over 217,000 men would be diagnosed with cancer of the prostate and about half of them were expected to have early stage, low-risk tumors.¹ Thus, management of early stage, low-risk prostate cancer is an important public health issue.

A well-defined clinical benefit of immediate therapy with curative intent has not been demonstrated unequivocally 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. However, both surgical and radiation treatments result in significant short- and long-term adverse events. Thus, all men who undergo treatment risk suffering the adverse events of the treatment, while only a portion of them may derive some benefit. The balance of potential benefits and harms may be different for different treatments, and such balances may vary by patient characteristics, including age, health status, and tumor characteristics.

Watchful waiting and active surveillance are two followup strategies for men with prostate cancer. The two terms are often used interchangeably in both the scientific community and the popular media. In this protocol, active surveillance means the monitoring of a potentially curable prostate cancer and intervening with a curative-intent treatment at the earliest sign of worrisome progression; watchful waiting means postponing therapeutic interventions until symptoms develop, with the primary objective being palliation of the symptoms rather than an attempt at a cure. Active surveillance is a fairly involved followup of patients—by monitoring PSA levels and sometimes also obtaining repeated prostate biopsies; whereas, watchful waiting is a relatively passive followup strategy, with interventions being triggered by symptoms.

The objective of this evidence review will be to summarize the existing literature on the role of active surveillance in the management of early stage, low-risk prostate cancer. Both the report and the corresponding NIH State-of-the-Science Conference are a part of the NIH Consensus Development Program. The purpose of the Program is to evaluate the scientific evidence on a particular topic and to develop a statement that advances research in this area. The statement is created by an independent panel that is assembled for the conference. The panel will hear the scientific data, including information from the evidence review, and then will use this information to compose their statement.

Additional information about the NIH Consensus Development Program can be found at: <http://consensus.nih.gov/>.

II. The Key Questions

The Key Questions (KQs) were provided by the conference sponsor. The KQs have not been altered except for style.

[AUTHORS—For accuracy in medical terminology, should “cores” be replaced with “core biopsies” below? Similarly, should “percentage of cores” be replaced with “percentage of positive core biopsies”?—EDITOR]

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?

- 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. Number of cores
 - iii. Histopathologic grading changes
- d. System Characteristics
 - i. Differences in geographical access

Question 2

How are active surveillance and other observational strategies defined?

- a. Common metrics



- i. Age
 - ii. Gleason Score
 - iii. Number of cores
 - iv. Percentage of cores
 - v. PSA (velocity, doubling time)
 - vi. Imaging
 - vii. Behavioral indicators
- b. Followup protocols
- i. Gleason Score
 - ii. Number of cores
 - iii. Percentage of cores
 - iv. PSA
 - v. Imaging
 - vi. Behavioral indicators

Question 3

What factors affect the offer of, acceptance, and adherence to active surveillance?

- a. Physician Factors
 - i. Primary care
 - ii. Diagnosing physician
 - iii. Consultant (second opinion)
 - iv. Clinical factors
- b. Patient Factors
 - i. Family involvement
 - ii. Personal preferences
 - iii. Risk perceptions
 - iv. Family history
 - v. Social support
- c. Delivery System Factors
 - i. Economic incentives and disincentives
 - (a). Insurance type (health maintenance organization, military, or private)
 - (b). Availability of technology
 - ii. Geographic location
 - (a). Small area variation

Source: www.effectivehealthcare.ahrq.gov

Published Online: May 19, 2011

- (b). Regional variation
 - (c). Urban versus rural
- iii. Academic centers versus private practice
- d. Communication Strategies
- i. Risk assessment; predictive models
 - ii. Decisionmaking tools and aids

Question 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

Question 5

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

III. Analytic Framework

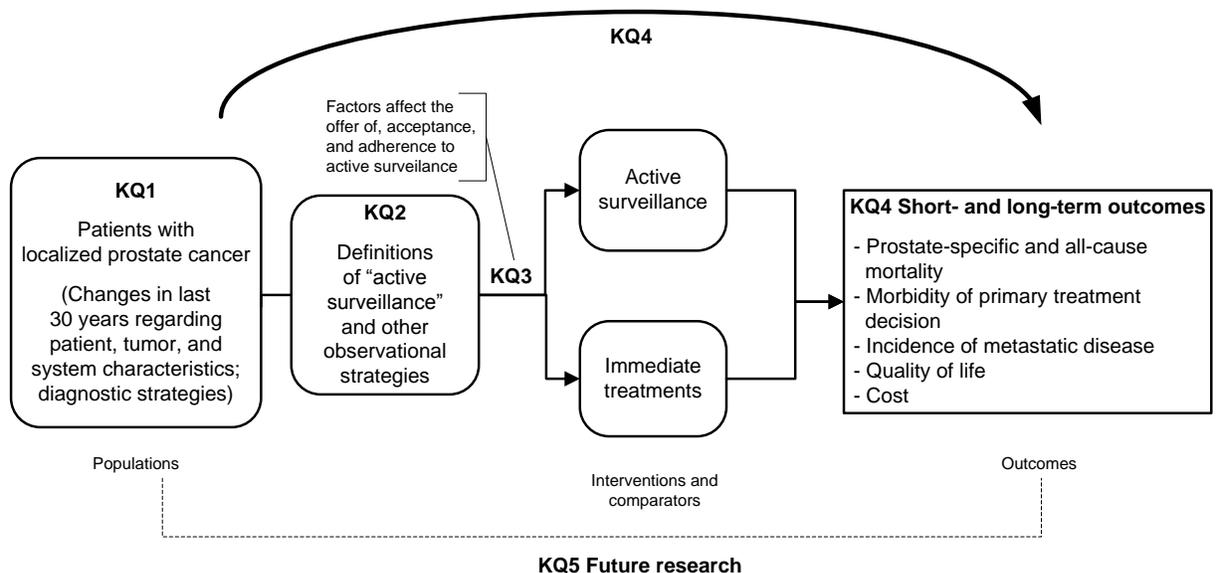


Figure. Analytic framework that depicts the five Key Questions (KQs) that examine the role of active surveillance in the management of men with clinically localized prostate cancer.

IV. Methods

A. Criteria for Inclusion/Exclusion of Studies in the Review

The study eligibility criteria are:

Population. Men with clinically localized prostate cancer (T1–T2), without (or unable to assess) either regional lymph node involvement (N0–X) or metastases (M0–X), regardless of age, histologic grade, Gleason Score, or PSA level.

Interventions. Active surveillance, watchful waiting, and other observational strategies.

Comparator. Immediate treatments, either curative or palliative.

Outcomes. Prostate-specific and all-cause mortality, incidence of metastatic disease, quality of life, morbidity from primary treatment, other adverse events, and costs.

Study design. Randomized controlled trials, nonrandomized comparative studies, and cohort studies; for adverse events, we will limit cohort studies to those with at least 100 patients; systematic reviews that used these eligibility criteria.

Based on the final inclusion and exclusion criteria, we will assess the titles and/or abstracts of citations identified from literature searches for inclusion. Full-text articles of potentially relevant abstracts will be retrieved for further examination. The reasons for excluding any full-text articles will be documented.

B. Literature Search

A comprehensive search of the scientific literature in Medline, the Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews will be conducted to identify studies of adult men relevant to each KQ. The search will be restricted to English-language publications due to time and resource restrictions. After discussion with the Technical Expert Panel (TEP) for the project (see section X below), it is agreed that the search dates should cover the last 30 years and will be from 1980 to the present with particular attention to the U.S. population. In addition to prostate neoplasms/cancer, search terms will include “active surveillance,” “watchful waiting,” “expectant management,” and others. In addition to citations from the literature search, we will include eligible publications suggested by the TEP. We will not include unpublished data sources or studies presented only in abstract form.

We will utilize existing systematic reviews as one component of the report to address the KQs. Systematic reviews will be selected based on direct relevance to the KQs with particular attention to the population, interventions, comparators, and outcomes covered. Depending on the findings of these reviews and the input received from the TEP, we will summarize literature published since the original search dates of these reviews.

Specifically, for KQ 1, we will use studies that analyzed the Surveillance, Epidemiology and End Results (SEER) database and other nationally representative cancer registries to address population and natural history trends, supplemented by studies included in relevant review articles and any additional eligible primary studies.

For KQ 2, the definitions of active surveillance and other observational strategies will be collected from all primary studies, as well as any existing systematic reviews and published guidelines/consensus statements that we identify. We will seek the TEP's help to establish study eligibility criteria.

For KQ 3, we will capitalize on existing systematic reviews and published guidelines/consensus statements on patient and physician factors affecting treatment decisions and on the type of communication strategies. The impact of insurance and geographic factors on therapeutic decisionmaking will be examined with targeted literature searches for primary studies and a recent comprehensive analysis of active surveillance.²

For KQ 4, we will rely on the multiple published systematic reviews that have compared the effectiveness of observational strategies (including active surveillance) with immediate treatment. These systematic reviews, collectively, analyzed a small number of randomized controlled trials and more than 100 observational studies.

For KQ5, future research needs will be informed by an evaluation of available evidence for KQs 1–4. There will not be a separate literature search for KQ 5.

C. Data Extraction and Data Management

Data from each study will be extracted by one experienced methodologist into standard forms in Microsoft Word. The basic elements and design of the forms will be the same as multiple forms we have used for other evidence reports and technology assessments. The elements will include study identifiers, study design characteristics, patient characteristics, tumor characteristics, diagnostic strategies, system characteristics, followup protocols, outcome definitions, results, and study quality. Prior to use, the form will be customized to capture all the relevant elements for the KQs. Outcome and predictor (e.g., patient characteristic) data to be extracted will be those elements explicitly stated in each of the KQs.

D. Assessment of Methodological Quality of Individual Studies

For those KQs with applicable studies, we will assess the methodological quality of those studies based on predefined criteria. For the assessment of systematic reviews, we will use the Assessment of Multiple Systematic Reviews (AMSTAR) instrument to evaluate the quality of the review.³ For the assessment of randomized and nonrandomized studies, we will adapt methods as described in the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (hereafter, *Methods Guide*).⁴ For randomized controlled trials, the *Methods Guide* mainly considers the methods used for randomization, allocation concealment, and blinding, as well as the use of intention-to-treat analysis, report of the study dropout rate, and the extent to which valid primary outcomes were described. For nonrandomized comparative studies, the *Methods Guide* suggests taking into account mainly the eligibility criteria, the similarity of the comparative groups in terms of baseline characteristics, methods for measuring exposure variables, and any adjustment for confounders.

We will apply a three-category quality-grading system (A, B, or C) to studies within each of the study designs to reflect the risk of bias (low, fair, or high).

E. Data Synthesis

We will summarize all the included studies in narrative form, as well as in summary tables that condense the important features of the study populations, design, intervention, and outcomes. Using the random-effects model, meta-analyses will be conducted when there are three or more unique studies with sufficiently similar study characteristics reporting the same outcome. The report will present findings, including the ordering of outcomes, in the order they appear in the KQs.

F. Grading the Strength of Evidence

When it is appropriate to do so, we will grade the strength of evidence for each of the outcomes addressed in the specific KQs by following guidelines described in the *Methods Guide*.⁵ We will determine the strength of evidence as high, medium, low, or insufficient to conclude based on four aspects: risk of bias, consistency, directness, and precision.

We will determine the risk of bias (low, medium, or high) based on the study design and the methodological quality of those studies.

We will determine the consistency of the data either as having no inconsistency or as having inconsistency present (or not applicable if only one study is available). We do not plan to use rigid counts of studies (e.g., 4 of 5 agree, therefore consistent), but instead we will evaluate the direction, magnitude, and statistical significance of all the studies and make a determination. We will describe our logic where studies are not unanimous.

We will assess the directness (direct or indirect) of the evidence. Indirect evidence will mean that either the studies used intermediate or surrogate outcomes instead of ultimate health outcomes or that the comparison of interest was not made in individual randomized controlled trials (e.g., that A vs. B can only be assessed by evaluating studies of A vs. placebo and B vs. placebo).

We will assess the precision (precise or imprecise) of the evidence based on the degree of certainty surrounding an effect estimate. A precise estimate is an estimate that would allow a conclusion to be drawn. An imprecise estimate is one for which the confidence interval is wide enough to include clinically distinct conclusions (e.g., both clinically important superiority and inferiority (i.e., the direction of effect is unknown), a circumstance that will preclude a conclusion.

G. Assessing Applicability

We will qualitatively assess the applicability of included studies to the U.S. population. This assessment will be based on the representativeness of United States-based studies (e.g., a national database vs. a single hospital) and on the similarity to the U.S. population of non-U.S. patients, prostate cancer screening standards, treatment, et cetera.

V. References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60:277-300.
2. Ollendorf DA, Hayes J, McMahon P, et al. Final appraisal document: active surveillance & radical prostatectomy for the management of low-risk, clinically-localized prostate cancer. Boston, MA: Institute for Clinical and Economic Review, 2009. Available at: <http://www.icer-review.org/index.php/Reports/asrp.html>. Accessed May 18, 2011.
3. 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.
4. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality, March 2011. AHRQ Publication No. 10(11)-EHC063-EF. Chapters available at: www.effectivehealthcare.ahrq.gov/methodsguide.cfm.
5. 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; 63:513-23.

VI. Definition of Terms

TNM = Tumor stage, node involvement, and metastasis.

VII. Summary of Protocol Amendments

In the event of protocol amendments, the date of each amendment will be accompanied by a description of the change and the rationale.

VIII. Review of Key Questions

For this State-of-the-Science Conference for the NIH Office of Medical Applications of Research (OMAR), the KQs are charged to OMAR, AHRQ, and the Evidence-based Practice Center (EPC) by the Conference Planning Committee and may not be changed. The KQs were reviewed with the TEP to assure that the review explicitly addresses the KQs.

IX. Key Informants

Key Informants are not applicable for this OMAR-sponsored review.

X. Technical Experts

Technical Experts comprise a multidisciplinary group of clinical, content, and methodologic experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted

opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the public review mechanism. In the case of reports generated for conferences sponsored by OMAR, the Conference Panel Chairperson will also serve as a member of the TEP.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for CERs and Technical briefs, be published three months after the publication of the Evidence report.

Potential Reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.