Contextual Review

Overview of Prostate Cancer Screening Decision Models: A Contextual Review for the U.S. Preventive Services Task Force

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. HHSA-290-2012-00015-I, Task Order No. 5

Prepared by:
Kaiser Permanente Research Affiliates Evidence-based Practice Center
Kaiser Permanente Center for Health Research
Portland, OR

University of Arizona
College of Medicine
Phoenix, AZ

Investigators:
Jennifer S. Lin, MD, MCR
Diana B. Petitti, MD, MPH
Brittany U. Burda, DHSc, MPH

AHRQ Publication No. 17-05229-EF-2
May 2018

This report is based on research conducted by the Kaiser Permanente Research Affiliates Evidence-based Practice Center (EPC) and the University of Arizona under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA-290-2012-00015-I, Task Order No. 5). The findings and conclusions in this document are those of the authors, who are responsible for its contents, and do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

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

Acknowledgments

The authors acknowledge the following individuals: David Meyers, MD, at AHRQ; current and former members of the U.S. Preventive Services Task Force; Ruth Etzioni, PhD, Roman Gulati, MS, David F. Penson, MD, MPH, Joshua J. Fenton, MD, MPH, and Joy Melnikow, MD, MPH, for providing expert review of the draft report; and Smyth Lai, MLS, and Elizabeth Hess, MA, ELS(D), at the Kaiser Permanente Center for Health Research.

Suggested Citation

Lin JS, Petitti DB, Burda BU. Overview of Prostate Cancer Screening Decision Models: A Contextual Review for the U.S. Preventive Services Task Force. AHRQ Publication No. 17-05229-EF-2. Rockville, MD: Agency for Healthcare Research and Quality; 2018.

Abstract

We reviewed a series of prostate cancer screening modeling studies most relevant to current U.S. practice to inform the U.S. Preventive Services Task Force on the magnitude of overdiagnosis using prostate-specific antigen screening, strategies to mitigate harms of screening, the dependence of model estimates of net benefit on values and/or preferences for various health states, and the net benefit of screening for populations at higher risk for prostate cancer or prostate cancer mortality. Modeling studies have the potential to inform our understanding of factors that influence the benefits and harms of screening. Comparative modeling studies with robust sensitivity analyses can provide estimates of upper and lower bounds of the benefits and harms of various screening strategies. Lowering the age to stop screening, lengthening the intervals between PSA testing, and raising the PSA threshold for biopsy can all mitigate harms of screening. Evidence that supports the choice of estimates of utility values for health states related to prostate cancer screening, diagnosis, and treatment is limited; estimates from existing studies vary widely for the same health state. For these reasons, the use of quality-adjusted life years as an integrated measure of the balance of benefits and harms of prostate-specific antigen screening is limited. To date, modeling studies that address screening in groups at higher risk for prostate cancer or prostate cancer mortality are conceptual in nature, such that studies that provide quantitative estimates of the balance of benefits and harms of prostate cancer screening in these subpopulations have not yet been conducted.

Chapter 1. Introduction

In 2012, the U.S. Preventive Services Task Force (USPSTF) recommended against prostate-specific antigen (PSA)–based screening for prostate cancer in men (D recommendation).[1] Based primarily on PSA screening trials, including the Prostate, Lung, Colorectal and Ovarian (PLCO) screening trial and the European Randomized Study of Screening for Prostate Cancer (ERSPC), the USPSTF concluded that the benefits of PSA-based screening for prostate cancer did not outweigh the harms. Based on these trials, the USPSTF determined that the reduction in prostate cancer mortality from screening was small and that these potential benefits were outweighed by harms from prostate biopsy and prostate cancer treatment of screen-detected tumors.

Immediately following the 2012 USPSTF recommendation, Cancer Intervention and Surveillance Modeling Network (CISNET) investigators Etzioni and colleagues at the Fred Hutchinson Cancer Research Center (FHCRC) wrote an editorial on the limitations of basing screening policies on screening trials alone.[2] They argued that model-based information should be considered alongside trial evidence. The USPSTF editorial that responded to specific criticisms acknowledged the limitations of trial data and issued counter-criticisms of model-based information, including concerns about the many assumptions required to construct models, problems with undetectable bias, lack of transparency, and issues regarding lack of replication.[3]

The controversy about whether to screen, or how best to screen, for prostate cancer is based on the assessment of the balance of benefits and harms. Because prostate cancer has (or can have) a long period when it is detectable but asymptomatic, harms resulting from overdiagnosis are particularly important to consider in understanding the balance of benefits and harms of screening for prostate cancer. Models that measure benefits and harms using an integrated measure (e.g., quality-adjusted life years[4]) have the potential to be helpful in assessing the benefits of screening (e.g., prostate cancer-specific mortality reduction) against the harms of diagnosis (e.g., biopsy) and subsequent treatment of screen-detected cancer (e.g., bowel, erectile, or urinary dysfunction).[5] Models could also identify screening strategies or critical factors in screening (e.g., age to start and stop screening, intervals of screening, criteria for biopsy or treatment) that may affect the balance of benefits and harms (e.g., by limiting overdiagnosis). Models could, therefore, provide information that would permit the design of strategies that maximize the potential benefits and minimize the potential harms of screening, and the selection of individuals in whom the balance of benefits and harms is more favorable, as well as provide information that could guide the management of individuals with screen-detected cancer. To support the USPSTF in updating its recommendation on screening for prostate cancer, we were asked to summarize the best available existing modeling studies to help frame the interpretation of the systematically reviewed empiric evidence on the benefits and harms of prostate cancer screening.[6]

Chapter 2. Methods

Questions

We worked with the USPSTF to identify important clinical questions for which models might provide information to help assess the balance of benefits and harms in prostate cancer screening. Our paper aims to review the model evidence most relevant to current U.S. practice in order to answer the following questions:

1. What do estimates derived from models suggest about the magnitude of overdiagnosis of prostate cancer due to PSA screening?
2. What do estimates derived from models suggest about the possible mitigation of harms of screening for prostate cancer by using different PSA screening strategies (i.e., varying age to start or stop screening, interval length between PSA testing, PSA thresholds to guide biopsy, and treatments)?
3. What is the dependence of model estimates of net benefit on values and/or preferences for various health states?
4. What do estimates derived from models suggest about the net benefit for subpopulations at higher risk for prostate cancer and/or prostate cancer specific mortality (e.g., African Americans [AA], persons with a family history of prostate cancer)?

Search

We conducted a targeted search for relevant articles in PubMed that were published from 2010 to April 6, 2016. We focused our database search to primarily capture modeling studies published after results were reported for PLCO and ERSPC. We used the following keywords: prostate cancer/neoplasms, PSA, mass screening, model, microsimulation, simulation, lifetime estimate, net benefit of screening, number needed to screen, ERSPC and PLCO. We identified additional citations from CISNET modelers as well as through searching the CISNET website (http://cisnet.cancer.gov/publications/cancer-site.html#Prostate) and reference lists of relevant articles. We identified 709 unique citations and subsequently reviewed 123 full-text articles.

Study Selection

Although we aimed to be as complete as possible in identifying relevant modeling studies, this is not a systematic review. Currently, robust methods to systematically review and critically appraise existing decision models do not exist. While we defined decision analytic models broadly— any health care evaluation model that accounts for events over time and across populations (microsimulation or otherwise), that is based on data drawn from primary and/or secondary sources (i.e., trials, observational data, life tables), whose purpose is to estimate the effects of PSA screening on important outcome measures of benefits and/or harms; we focus our analysis on the most relevant models based on assessment and demonstration of their validity, the transparency of the assumptions made, and the applicability of the models to current practice

in the United States. At times, we discuss older literature for historical context, but generally our analyses focused primarily on the major modeling studies that have been published since the completion of the large PSA screening trials PLCO and ERSPC. We included eight unique models in 29 publications (**Table 1**). Several identified modeling studies were excluded from our discussion because they were obsolete (i.e., no longer applied to current clinical practice), simplistic, and/or poorly described.[7-19] We did not address pre-biopsy prostate cancer risk prediction tools/calculators or adjunctive testing performed with PSA screening and/or prostate biopsy (e.g., Prostate Health Index, *PCA3*) to mitigate harms, as these are covered in the accompanying systematic review of empiric evidence on the benefits and harms of prostate cancer screening.[6] Several publications that contain important contextual information on overdiagnosis or utilities are referenced but not included in the **Table 1** because they are not modeling studies.

Data Analyses

Our analyses of the model data primarily focuses on the CISNET consortium of modeling studies. The most recent publications from the CISNET prostate cancer modelers use the following names and acronyms to describe their models: the Erasmus MIcrosimulation SCreening ANalysis Prostate Cancer (MISCAN-PRO), the FHCRC Prostate-Specific Antigen growth and Prostate Cancer progression (PSAPC), and the University of Michigan Self-Consistency Analysis of Surveillance (SCANS) (**Table 2**). The FHCRC model, previously called Prostate Cancer SIMulation (PCSIM), was renamed PSAPC in 2009, after a new parametrization was incorporated to facilitate empirical estimation of key relationships, including relationships between PSA growth and cancer progression.

While we do not formally critically appraise the included model studies, we attempt to identify the most important assumptions and limitations to our understanding of the included models' estimates and findings. We summarize the included modeling studies' findings in context these assumptions and limitations for each of the four questions (overdiagnosis, strategies to mitigate harms, utilities, and subpopulations). We limit our analyses primarily to modeling studies using U.S. data; however, when appropriate we also discuss key modeling papers using data derived from other countries. We generally discuss modeling studies chronologically as many of the models build on previous work. Given the importance and complexity of the issue of overdiagnosis in screening for prostate cancer, this paper is accompanied by a more detailed analysis that used model data to understand overdiagnosis.[20] Several modeling studies included cost or cost-related outcomes (e.g., willingness to pay), which we do not discuss.

Expert Review and Public Comment

A draft of this report was shared with invited expert reviewers. We compiled and addressed (where appropriate) the comments received from these invited experts. Additionally, a draft of the full report was posted on the USPSTF Web site from April 11, 2017 through May 9, 2017. A few comments were received during this public comment period; no changes were made to the report based on these comments.

USPSTF Involvement

We worked with several USPSTF members to determine the questions and scope for this report. AHRQ funded this work under a contract to support the work of the USPSTF. An AHRQ Medical Officer provided project oversight, reviewed the draft report, and assisted in the external review of the report.

Chapter 3. Results

Overview

We found a number of relevant modeling studies that addressed the issue of overdiagnosis in PSA screening (k=22), strategies to mitigate harms of PSA screening (including but not limited to overdiagnosis) (k=11), the dependence of model findings on patient values and (dis)utilities (k=7), and screening in subpopulations at higher risk (based on risk factors other than age) for prostate cancer and/or cancer mortality (k=3) (**Table 1**). In addition to the three CISNET models (MISCAN-PRO, PSAPC, and SCANS), we include five additional unique models that address prostate cancer screening. These models vary in complexity and transparency. All but one[21] can be broadly described as state transition models, such that the cancer process is conceptualized as a series of events from no cancer through cancer and death occurring over time.

To date, the CISNET models of prostate cancer screening represent the best models to evaluate PSA screening effects on patient outcomes. All three CISNET prostate cancer models are quite complex and have many assumptions (**Table 2**). In addition, the models and their assumptions have evolved in efforts to improve their ability to address new questions, to incorporate better empiric information, and to make them more representative of contemporary experience. The PSAPC and MISCAN-PRO models are microsimulation models, defined by CISNET as "computer models that operate at the level of individuals or smaller entities such as tumors or cells." Both the PSAPC and the MISCAN-PRO models use algorithms and random draws from parametric statistical distributions to estimate natural history and outcomes. In contrast, the SCANS model represents the cancer process in terms of a series of equations that have a closed-form solution; the outcomes are derived analytically or numerically.

These complex models must make assumptions, at least in part to avoid potential biases of simpler models. We describe a few important assumptions of (but not limited to) the CISNET models here. All models make assumptions about the natural history of prostate cancer (e.g., how progression depends on age, PSA, or stage/grade of tumor). When used to provide data pertinent to the United States, the models are calibrated to U.S. Surveillance, Epidemiology, and End Results (SEER) program data, which are the best available data but are not wholly representative of the United States. All three CISNET models assume that all cancers progress over time and that all would eventually become metastatic. The three CISNET models differ in their assumptions about the length of the preclinical detectable phase, how much early detection improves tumor characteristics, and how natural history and screening depend on age. Unlike the MISCAN-PRO or SCANS models, the PSAPC model models PSA growth (i.e., generates a PSA level for each individual at each screening); the natural history of prostate cancer (onset, progression to clinical diagnosis, progression to metastasis) is dependent on age and PSA level. Second, the models differ in the way they conceptualize how screening might generate a survival benefit. For the PSAPC and SCANS models, the benefit of screening arises from a shift to an earlier stage of cancer (stage shift) due to screening. For the MISCAN-PRO model, the benefit arises due to the cure of cancers detected by screening; in this model, the mortality benefit due to screening is calibrated to the ERSPC trial. Models that attempt to assess the effect of screening on survival make assumptions about the baseline incidence of prostate cancer and survival in the

absence of screening. CISNET models use historical SEER data prior to the advent of screening to estimate prostate cancer survival following clinical diagnosis. This approach is not ideal, but a better alternative appears not to exist. Finally, models make assumptions about the distribution of treatment (according to age, stage, and grade) and about the efficacy of prostatectomy, radiation, hormone therapy, and conservative management (i.e., watchful waiting or active surveillance) in screen-detected cancer. The CISNET models estimate the treatment benefit for localized cancer using data from the Scandinavian Prostate Cancer Group (SPCG-4) trial, which was not conducted in men with screen-detected cancers.

Question 1. What Do Estimates Derived From Models Suggest About the Magnitude of Overdiagnosis of Prostate Cancer Due to PSA Screening?

We identified 22 modeling studies that examined the magnitude of overdiagnosis of prostate cancer due to screening (**Table 1**). In our summary of the most relevant data, we focus on a subset of these studies.[21-27] Additional studies and details on how to consider model data in order to understand overdiagnosis and other questions about overdiagnosis addressed in modeling studies are described in a companion paper.[20]

In the literature on prostate cancer modeling, overdiagnosis is almost always defined as screen-detected cancer that would not have been clinically detected during a patient's lifetime in the absence of screening; the number of overdiagnosed cancers is the numerator used in measures of the "frequency" of overdiagnosis. However, studies use different denominators to measure overdiagnosis: overdiagnosed cancers as a proportion of 1) the total number of screen-detected prostate cancers during a specified period of time, 2) the total number of prostate cancers (i.e., both screen-detected and clinically detected) during a specified period of time, and 3) the total number of men screened (or eligible to be screened) at the starting point for screening. Estimates of overdiagnosis in the modeling studies use the lead time approach. The model is used to derive estimates of lead time (i.e., time between detection by screening and when the cancer would have been clinically detected). If the lead time of the cancer exceeds the life expectancy of the individual being screened, then the cancer is overdiagnosed (i.e., the man would have died before the clinical detection of the prostate cancer).

Early modeling studies to evaluate overdiagnosis demonstrated that PSA screening advances the diagnosis of prostate cancer in time and is associated with overdiagnosis[22, 24] and demonstrated that age at screening and sojourn time (i.e., the time between initiation of the cancer and when it would be clinically detected) are important factors influencing the rates of overdiagnosis.[25] Draisma and colleagues[22, 23] used the MISCAN-PRO model with data from the Rotterdam section of the ERSPC to estimate how much overdiagnosis there might be in the Netherlands for nine hypothetical screening strategies based on a PSA test threshold for biopsy referral of greater than 3 micrograms per liter (µg/L). The range of estimates of overdiagnosis (as a proportion of screen-detected cancers) ranged from 27% for a single screen at age 55 years to 56% for annual screening at 55 to 75 years or a single screen at age 75 years. Using a basic analytic modeling approach, Davidov and Zelen[25] explored a range of hypothetical values of sojourn time (from 5 to 20 years) for various hypothetical screening schedules and test sensitivities and concluded that

the frequency of overdiagnosis (as a proportion of screen-detected cancer) was in the range of 20% to 40% for most "realistic" values of sojourn time and ages of screening.

Modeling studies subsequently attempted to estimate how much of the observed increase in the incidence of prostate cancer after introduction of PSA screening in the United States for the period from the middle to late 1980's (when PSA screening began to become widespread) through 2000 might be overdiagnosed.[21, 26, 27] In the most comprehensive of these modeling studies, Draisma and colleagues[27] compared estimates of overdiagnosis based on three CISNET models for men screened from 1985 to 2000 in the United States. All three models were calibrated to U.S. SEER data and used U.S. data on life expectancy. The proportion of all screen-detected prostate cancer in the United States during 1988 to 2000 that was overdiagnosed was estimated to be 22.9% (SCANS), 28.0% (PSAPC), and 42.0% (MISCAN-PRO); the corresponding proportion of all cases of prostate cancer in the United States that were overdiagnosed was estimated to be 8.6% (SCANS), 11.9% (PSAPC), and 18.6% (MISCAN-PRO).

A number of modeling studies directly address how different screening strategies affect overdiagnoses. These studies are addressed in the following section on strategies to mitigate harms (Question 2).

Summary

In evaluating estimates of overdiagnosis from prostate cancer screening models (as well as empiric studies), it is important to recognize that studies may use different definitions of overdiagnosis and sometimes the specific metric being used is difficult to assess, making comparisons of estimates of overdiagnosis difficult. Estimates of overdiagnosis are dependent on many factors that can vary across populations (e.g., natural history of cancer, life expectancy of the population, prior history of screening). As such, to understand estimates of overdiagnosis, it is important to be clear about the metric being used, what population is being addressed, and the assumptions and limitations of the model to characterize both the natural history of the cancer and the population being studied. From modeling studies using U.S. data, it is reasonable to conclude that the proportion of screen-detected prostate cancers that were overdiagnosed in the United States during the late 1980's through 2000 is substantive, with best estimates between 23% and 42%.[27] The magnitude of overdiagnosis from prostate cancer screening will necessarily depend on the screening strategy (i.e., age to start and stop screening, intervals, PSA threshold for biopsy) and subsequent diagnostic follow-up (e.g., adherence to biopsy, number of biopsy cores). How different screening strategies can mitigate harms, including overdiagnosis, is discussed next.

Question 2. What Do Estimates Derived From Models Suggest About the Possible Mitigation of Harms of Screening for Prostate Cancer by Using Different PSA Screening Strategies?

We included 12 modeling studies that directly addressed the effect of different screening strategies to mitigate harms, eight of which explicitly address overdiagnosis as a harm (**Table 1**). Here, we focus on the most recent relevant modeling studies with U.S.-specific data, which includes both the PSAPC model[28, 29] and the MISCAN-PRO model.[30] Additional details, including a discussion of studies addressing overdiagnosis with non-U.S. data, are in our companion paper.[20]

A 2013 publication by Gulati and colleagues[28] used the PSAPC model to evaluate the comparative effectiveness of 35 screening strategies (i.e., varying age to start and stop screening, screening intervals, and PSA thresholds for biopsy) using U.S. data. The model-projected age-adjusted incidence of prostate cancer closely matches the observed incidence through 2005. In addition, a model simulation of the ERSPC projected a 28% reduction in prostate cancer mortality after 11 years of follow-up, which is close to the 29% reduction observed in this trial after correction for non-adherence. The study examined a number of outcomes, including the probability of prostate cancer death, the probability of life saved, and the meantime of live saved (in months) to measure benefit, the probability of one or more false-positive results, and the lifetime probability overdiagnosis (as a proportion of individuals screened) to measure harm. In addition to reporting these measures of benefit and harm for each of the 35 screening strategies, the modelers used the metric "additional number needed to detect (NND) to prevent one prostate cancer death" as a summary measure of the trade-off of harms to benefits. For the reference strategy (**Table 3**, strategy 8), annual screening for men 50 to 74 years with a PSA threshold for biopsy of 4.0 µg/L, the model estimated a lifetime probability of overdiagnosis of 3.3%, a probability of prostate cancer death of 2.15%, and a probability of life saved of 0.70%. This was a relative reduction of 24.8% in the probability of prostate cancer deaths compared with the 2.86% probability of prostate cancer death with no screening. The estimates of the lifetime probability of overdiagnosis ranged from as high as 6.0% for the complex National Comprehensive Cancer Network (NCCN) strategy (strategy 1, interval of screening varies by age and PSA level, biopsy threshold varies by PSA level or velocity) with men aged 40 to 74 years to as low as 1.3% for biennial screening of men aged 50 to 69 years with a PSA threshold of 4.0 µg/L. Corresponding probabilities of prostate cancer death ranged from 2.02% for the NCCN strategy to 2.43% for biennial screening of men aged 50 to 69 years with a PSA threshold of 4.0 µg/L (**Table 3**). Thus, aggressive screening strategies (particularly with low PSA thresholds for biopsy referral) optimize benefit (i.e., prostate cancer deaths prevented, lives saved) but generate substantial harms (i.e., overdiagnosis). Using the NND to understand the trade-off between overdiagnosis and probability of life saved demonstrates that less frequent testing (strategy 22, screen at ages 45 to 74 years biennially or quinquennially if PSA level is less than the median for age) and a more conservative threshold for biopsy in older men (strategy 20—screen at ages 50 to 74 years annually with PSA level greater than the 95th percentile for age as threshold for biopsy referral) can preserve much of the survival effect and reduce the harms of screening compared to the reference strategy (**Figure 1**). For strategy 22, the estimated lifetime probability

of overdiagnosis is 2.4% and the estimated probability of a life saved is 0.58%. For strategy 20, the estimated lifetime probability of overdiagnosis is 2.3% and the estimated probability of a life saved is 0.61%.

de Carvalho and colleagues[30] conducted a similar modeling exercise with the MISCAN-PRO models that used U.S. data to evaluate 83 strategies varying by age to start and stop screening, screening intervals, and PSA thresholds for biopsy. Similarly, they presented the benefit using the probability of prostate cancer death and harm using the probability of overdiagnosis (as a proportion of individuals screened). Although the hypothetical screening strategies are different from those evaluated in the previously described PSAPC study, the model findings are consistent with the study by Gulati and colleagues.[28] For the reference screening strategy (annual screening for men aged 50 to 74 years, PSA threshold of 3.0 µg/L [as opposed to 4.0 µg/L in the previously described PSAPC study]), the estimated lifetime probability of overdiagnosis was 3.8% and the corresponding estimate of prostate cancer death was 2.38% (**Appendix Table 1**). The lowest estimate of the lifetime risk of overdiagnosis (overdiagnosis as a proportion of individuals screened) was 1.99% for a strategy in which men aged 50 to 70 years are screened every 4 years with a PSA threshold for biopsy of 3.0 µg/L. The corresponding probability of prostate cancer morality for this strategy was 2.53%. Among strategies with similar benefit in reducing prostate cancer mortality, using a lower age to stop screening reduces overdiagnosis more than lengthening the screening interval (i.e., yearly up to every 4 years) or raising the PSA threshold (i.e., from 3.0 µg/L to 4.0 µg/L) (**Appendix Table 1**).

A 2016 publication by Roth and colleagues[29] that used the PSAPC model evaluated the net benefit of screening using estimated life-years (LYs) and QALYs for 18 strategies for screening and treatment (**Table 4**) informed by the comparative effectiveness analysis by Gulati and colleagues[28] discussed previously. The strategies presented were chosen to reduce the harms (i.e., overdiagnosis) by more than half (compared with the reference strategy) while maintaining most of the benefits (i.e., lives saved). This modeling study examined 18 possible screening strategies that varied by age to start and stop screening, screening intervals, and tailored criteria for biopsy referral. This study also evaluated the possible use of more selective treatment and conservative management (active surveillance) for low-risk screen-detected prostate cancer cases as strategy to reduce harm. Earlier modeling studies have suggested that active surveillance, as opposed to immediate prostatectomy, for low-risk disease (i.e., Gleason grade 6, stage ≤T2a) permits the avoidance of a substantial proportion of surgeries (overtreatment) without a substantial increase in prostate cancer mortality.[15, 16, 31] However, the ProtecT trial, published in 2016, reported that men with screen-detected localized prostate cancer randomized to active surveillance had an increased risk of metastatic prostate cancer compared with men randomized to immediate radical therapy (radical prostatectomy or radical radiotherapy) at a median of 10 years after randomization, with no differences in prostate cancer mortality between the three groups.[32]

Eight of the 18 screening strategies modeled by Roth and colleagues[29] evaluated possible use of a PSA threshold for biopsy referral of 10.0 µg/L. For the 10 screening strategies that evaluated PSA threshold for biopsy other than 10.0 µg/L, the estimated effect on LYs and QALYs of a selective treatment strategy (i.e., active surveillance instead of immediate treatment for low risk cancer) was evaluated. The authors noted that active surveillance as used in the modeling exercise (i.e., curative treatment not offered until cases progress to "would-be clinical

diagnosis") is not consistent with contemporary practice in which curative treatment would likely be offered at an earlier point.[29] The model examined both LYs gained and QALYs. With the exception of the healthy state, which was assumed to have a utility of 1.0, the health state utilities used to estimate QALYs were derived from a single study.[33] The limitations of the model's use of utilities are discussed in the section on values and/or preferences for various health states (Question 3). This modeling study suggested that only the strategies with a biopsy threshold at a PSA level of 10.0 µg/L and strategies that incorporated selective treatment practices (which were considered only for the 10 strategies with a PSA threshold less than 10.0 µg/L) increased QALYs compared to no screening and that the absolute magnitude of QALYs gained for these strategies was small (0.002 to 0.005) (**Table 4**).

Other Relevant Information From Modeling Studies

We also discuss an additional important modeling study brought to our attention during the drafting of this report. This study by Tsodikov and colleagues[34] used individual-level data from both the ERSPC and PLCO trials in collaboration with the three CISNET prostate cancer screening modeling groups to evaluate whether screening efficacy differed between the ERSPC and PLCO trials after accounting for differences in screening and diagnostic follow-up across trial arms. While this modeling study did not meet relevance for inclusion on strategies to mitigate harms, it does suggest that the apparent differences in screening benefit comparing the "intention to treat" analyses in the PLCO and ERSPC trials can be explained by differences in the "intensity" of screening and subsequent diagnostic follow-up (e.g., PSA threshold, frequency of screening, biopsy referral uptake). To quantify the "intensity" of screening, the modelers estimate what they term "mean restricted lead time "(MRLT), which is a version of mean lead time (extent to which diagnosis is advanced under screening). This modeling study suggested that 1) screening reduced the risk of prostate cancer mortality by 16% (95% CI 4, 27%) after accounting for different baseline risk of prostate cancer death in the PLCO relative to the ERSPC and age at randomization, and 2) screening was associated with a 7 to 9% lower risk of prostate cancer mortality per year of MRLT, or 6 to 15%, 22 to 28%, 28 to 35%, and 31 to 37% reductions in expected prostate cancer mortality in the ERSPC control, PLCO control, ERSPC screening, and PLCO screening arms, respectively, over 11 years of follow-up compared to no screening. These findings are concordant with the findings of modeling studies previously discussed which demonstrate that more intensive screening has greater mortality benefit. However, this mortality benefit must be weighed against potential harms of more intensive screening for prostate cancer. Data on estimated harms were not presented.

Summary

Strategies to mitigate harms of screening include decreasing the age to stop screening, lengthening intervals between screening, and raising the PSA threshold for biopsy and strategies that would implement selective treatment. Ranking strategies by net benefit is not easy because maximizing benefit (i.e., reducing prostate cancer mortality) necessarily increases harm (overdiagnosis) and because the weighing of benefits to harms is subjective. For example, there is no consensus on how many overdiagnosed cases of prostate cancer would be acceptable to prevent one prostate cancer death. Modeling using QALYs as an outcome attempts to account

for both benefit and harm using a single measure. The most recent modeling study using QALYs demonstrates that only strategies with a biopsy threshold at a PSA level higher than 10.0 µg/L, or those strategies with selective treatment practices (active surveillance) increased QALYs compared to no screening, and the incremental gain in QALYs was very small. The current use of QALYs in prostate cancer screening modeling to weigh benefits and harms has important limitations, which we discuss next.

Question 3. What Is the Dependence of Model Estimates of Net Benefit on Values and/or Preferences for Various Health States?

We identified three prostate cancer models in seven publications that incorporated or examined how values and/or preferences for various health states (i.e., utilities) affect estimates of the net benefit of prostate cancer screening.[29, 35-40] One publication reported on QALY using the PSAPC model,[29] three using the MISCAN-PRO model[35, 36, 40] and another three using a model from North Carolina State University (NCSU).[37-39] An early publication by Krahn and colleagues[9] that used a model to evaluate the effect of prostate cancer screening on quality of life is not discussed further because it evaluates screening strategies no longer used in clinical practice and because it derived utility estimates from a very small number of physicians. This method for eliciting utilities is now considered obsolete.

The 2016 PSAPC modeling study by Roth and colleagues,[29] which used U.S. data described previously, relied on a single study by Stewart and colleagues[33] to derive all of the estimates of disutility for the six health states associated with cancer screening and treatment that was considered: symptomatic with prostate cancer, surveillance for prostate cancer, short-term treatment for prostate cancer, long-term treatment for prostate cancer, distant stage prostate cancer, and end of life due to prostate cancer (**Appendix Table 2**). The model did not incorporate the (dis)utility of prostate cancer biopsy or having an elevated PSA below the threshold for biopsy (10.0 µg/L). The study by Stewart and colleagues,[33] which is the basis for the utilities used by Roth and colleagues,[29] elicited preferences for 19 prostate cancer health states from men age 60 years and older (n=162) living in the San Diego area, 52% of whom had been diagnosed with prostate cancer.[33] To reduce the response burden, the study had subjects assess nine of the 19 health states by using the standard gamble, time trade-off, and visual analog scale methods except that all subjects were asked to assess impotence and bowel problems using all three methods. Detailed data on elicited utilities are presented only for the standard gamble method; these standard gamble estimates were used in the modeling study by Roth and colleagues.[29] How the information on utilities for 19 health states presented by Stewart and colleagues was used to estimate decrements in utilities for the six health states considered in Roth and colleagues'[29] modeling study was not described in detail. In the study by Stewart and colleagues, [33] the range of utilities for each of the 19 health states was generally large; for many important health states (impotence, bowel problems, impotence and bowel problems, radiation therapy, and prostatectomy), the range of elicited standard gamble utilities was 0.0 to 1.0. Even the interquartile range was large for important utilities—0.01 to 0.52 for metastatic cancer, 0.56 to 0.90 for prostatectomy. The number of men providing information for the health states ranged from 38 (orchiectomy) to 150 (impotence) and 152 (bowel problems). Stewart and colleagues[33]

reported that age was a significant predictor of higher utility ratings for urinary difficulty and lower ratings for bowel problems. Utilities for impotence were slightly higher among those who had experienced impotence, and experience with urinary incontinence also increased ratings for this health state.[33] Overall, in the analysis by Roth and colleagues,[29] incorporating information on utilities (QALY) decreased the magnitude of the estimated benefit of screening (LY) by an order of magnitude (e.g., from 0.04 LY gained to 0.004 QALY gained compared to no screening). For the strategies with a PSA threshold for biopsy less than 10.0 µg/L, incorporating information on utilities changed the magnitude of the estimated benefit using LY to a negative (net harm) using QALYs. Roth and colleagues[29] stated that one-way sensitivity analyses on QALYs demonstrated that "results were by far most sensitive to the health states utility in the conservative management state." However, detailed data from the one-way sensitivity analyses was not presented. One-way sensitivity analyses that vary the utility values to reflect the range of values reported in the primary studies used to select utilities for base case analyses or values chosen for plausibility are essential to helping us understand the robustness and certainty of findings using QALYs. Two-way and three-way sensitivity analyses that vary utility values for health states found to be sensitive in the one-way sensitivity analyses would also be useful.

Three publications based on the MISCAN-PRO model evaluated the benefit of prostate cancer screening and health-related quality of life.[35, 36, 40] Two publications assessed various screening strategies (varying age, interval, and number of screenings) using a number of outcomes, including relative reduction in prostate cancer-specific mortality, LYs, and QALYs.[36, 40] Both publications used the same model and model inputs, including estimates for utilities of various health states. The 2012 publication is the primary modeling study[36]; subsequent analyses for cost-effectiveness using QALYs gained were provided in the 2015 publication.[40] As we do not discuss cost or cost-related outcomes, we focus on the 2012 publication here. In this study, estimates for the utilities of various health states are derived from a number of sources, including the ERSPC trial, cost-effective analyses for various prostate cancer treatments, and primary studies eliciting utilities from patients (**Appendix Table 3**). The authors used health-related quality-of-life data from the ERSPC population in Rotterdam to estimate the utility for screening itself (0.99 as screening has little effect on short-term health status and anxiety). The disutility of prostate biopsy was extrapolated from breast-cancer specific studies. This model, which used primarily ERSPC data, modeled the effects of prostatectomy, radiation therapy, and active surveillance. This model incorporated two studies of quality of life after prostate cancer treatment from men participating in the ERSPC trial in Rotterdam and Gothenburg, which found that after treatment among men without previous symptoms, 83% to 88% of men undergoing radical prostatectomy and 42% to 66% of men undergoing radiation therapy had erectile dysfunction; and 6% to 16% of men undergoing radical prostatectomy and 1% to 3% of men undergoing radiation therapy had urinary incontinence. In the base case, annual screening of men age 55 to 69 years (with a PSA threshold of 3.0 µg/L [range, 2.5 to 4.0 µg/L] for biopsy), the model estimated a 73 LY gain per 1,000 men compared to no screening (**Table 5**). Taking into account the adverse effect on quality of life (due to 247 additional negative biopsies and 41 additional men receiving prostate cancer treatment), the model estimated 56 QALYs gained per 1,000 men. In sensitivity analyses that varied assumptions about overdiagnosis, rates of attendance, and the values of the utilities (for screening, biopsy, cancer diagnosis, radiation therapy, prostatectomy, active surveillance, post-recovery period, palliative therapy, and terminal illness) (**Appendix Table 3**), the estimates of benefits for the base-case screening strategy

(annual screening of men age 55 to 69 years with a PSA cutoff value of 3.0 µg/L) varied from 97 QALYs gained per 1,000 men using the most favorable utility estimates to 21 QALYs lost per 1,000 men using the least favorable estimates. A substantial part of the estimated difference between LYs and QALYs gained was due to overdiagnosed cases (estimated as 42% of cancers detected during screening). Therefore, strategies to reduce overdiagnosis (and subsequent overtreatment) has the potential to increase the QALYs gained. In addition, the utility for the post-treatment period had a considerable effect on estimates of QALYs.

The other MISCAN-PRO paper aimed to describe a utility "break-even point," that is, the value of the utility (or utility level) below which the expected loss in quality of life due to earlier diagnosis and treatment of prostate cancer exceeds the expected benefits from preventing prostate cancer mortality.[35] This information might help men decide whether to undergo prostate cancer screening based on a personal judgement about the amount of loss in quality of life they would be willing to give up because of screening (and treatment) compared with the potential gain from avoided death due to prostate cancer. This model estimated that the utility break-even points were high (0.947 to 0.960) for screening at all ages, suggesting that screening should be avoided in men who judge that their quality of life would be decreased by more than 4% to 5% after cancer diagnosis and treatment. In sensitivity analyses that varied assumptions about lead-time, incidence, survival rate, and cure rate, the estimates of break-even points varied from 0.833 to 0.991. As expected, reduction in overdiagnosis (shorter lead times), lower survival rates after diagnosis, and higher cure rates due to treatment yielded results more favorable to screening. This work highlights the conceptual importance of a break-even point that is specific for an individual. The information might be useful in designing a decision aid. The authors acknowledged that expected utility theory is not always a good predictor of a patient's actual decisions.

Lastly, two publications using the same model, a seven-stage Markov model using U.S. data, compared different screening strategies and used QALYs.[37, 38] This model is poorly described and used only three health states (biopsy, treatment with prostatectomy, and metastatic disease). The disutilities for treatment and metastatic disease are based on a systematic review of the literature on utilities by Bremner and colleagues[41]; the utility for prostate biopsy is based on extrapolation from breast and bladder cancer studies. The review by Bremner and colleagues[41] of utilities for prostate cancer health states included 23 articles and presented data on 173 unique utilities; most studies had small sample sizes. Despite the numerous limitations of this model (including the assumption that all patients with prostate cancer undergo prostatectomy) and model inputs regarding utilities, this model confirms or supports the finding that the comparative performance of screening strategies are highly dependent on disutility for metastatic prostate cancer, prostate cancer treatment and, to a lesser extent, prostate cancer biopsy.

Summary

In modeling studies that incorporated estimates of utilities for health states after screening and treatment, the estimates of the magnitude of the benefit of screening programs considering QALYs compared with benefits measured using LYs was substantially reduced and, for some screening scenarios, became negative (indicating net harm). Overall, the inputs for the utilities used in these modeling studies were drawn from a fairly selective body of evidence that often

included small samples and found a wide range in the value of utilities for the same health states; and sensitivity analyses from modeling studies demonstrated that the net benefit using QALYs is dependent on the estimates of various utilities for health states. In fact, Gulati and colleagues[28] at FHCRC noted in their 2013 publication of the comparative effectiveness of different PSA screening strategies that the data informing utilities related to prostate cancer screening and post-diagnosis health states was extremely limited and not reliable for modeling. Therefore, using QALYs an outcome measure demonstrates that there is still considerable uncertainty about the net benefit of prostate cancer screening on a population level, and considerable variation, at the individual level, of preferences for various health states.

Question 4. What Do Estimates Derived From Models Suggest About the Net Benefit for Subpopulations at Higher Risk for Prostate Cancer and/or Prostate Cancer-Specific Mortality?

Race or ethnicity and family history (including but not limited to inherited genetic mutations) are two well described risk factors for developing and/or dying from prostate cancer. As of 2016, we were able to identify only three modeling studies that examined the impact of risk-stratified PSA screening in specific populations at higher risk for prostate cancer.[42-44] Only one of these modeling studies assessed possible differences in the balance of benefits and harms for prostate cancer screening in African American (AA) men compared with the general population.[44] A 1994 modeling study by Krahn and colleagues separately estimated the effect of prostate cancer screening on QALYs for general and AA populations; however; we do not discuss it further because of the limitations we previously described.[9]

One modeling study by Yen and colleagues is a six-state Markov model based on data from the Finnish trial within the ERSPC.[42] This was a simplistic model of the natural history of prostate cancer and did not explicitly include biopsy, diagnosis, and treatment of prostate cancer. The authors themselves questioned the ability of their model's natural history parameters to adequately capture progression from the preclinical detectable phase to clinical disease (sojourn time) and incorporate the (presumed) effect of genetic variants. Notwithstanding the acknowledged limitations, the authors presented the model-based evaluation of a risk-stratified screening protocol (varying ages and intervals) for three prostate cancer-related single-nucleotide polymorphisms (SNPs; rs4242382, rs200331695, rs138213197), which were chosen based on a review of the literature. The frequencies and natural history of prostate cancer for these SNP variants were derived from a Finnish study of hereditary factors in prostate cancer.[45, 46] Based on odds ratios quantifying the association of the SNPs with the risk for prostate cancer, two initiators were assumed to affect the onset of prostate cancer (rs138213197, rs4242382) and one promoter was assumed to affect the progression to an aggressive cancer (rs200331695). This modeling study was a proof-of-concept study meant to demonstrate that starting screening at an earlier age might be more beneficial in subpopulations with a genetic predisposition to developing prostate cancer, while shorter screening intervals might reduce interval cancers in subpopulations with a predisposition to more aggressive cancers (and thus help personalize surveillance strategies or treatments for prostate cancer).

Another recent modeling study using the PSAPC model[43] examined the implications of PSA screening in U.S. men with germline *BRCA* mutations, based on data from the IMPACT study (Identification of Men with a genetic predisposition to ProstAte Cancer: Targeted screening in *BRCA1/2* mutation carriers and controls).[47] It is known that men who are *BRCA1/2* mutation carriers have both a higher incidence of prostate cancer, an increased risk of metastasis, and a worse survival rate compared with non-carriers. The primary aim of the modeling study was to develop a conceptual framework for determining how to screen men with an increased risk for prostate cancer rather than actual guidance or recommendations for risk-stratified screening. The authors acknowledged that the model was a simplification of PSA growth and cancer progression and may not generalize well to subpopulations due to the complexities of biological processes as well as interventions. In addition to the inherent limitations and assumptions in the model as previously described, the IMPACT study offers limited data, with selectively recruited participants, only one round of screening, and minimal information on prior PSA testing. The limitations of the model, the data from IMPACT, and other important assumptions (e.g., biopsy frequency and sensitivity are the same for *BRCA1/2* mutation carriers as average risk men, treatment benefit is the same for carriers as for average-risk men, and the increased risk of metastases is constant) reinforce that this exercise was conceptual in nature. To model prostate cancer outcomes in men with increased risk, the model investigated a range of multiplicative factor(s) to the risk of disease onset, progression to symptomatic or metastatic state, and/or progression to death due to prostate cancer. However, it is unclear if this study conducted sensitivity analyses to determine the robustness of these assumptions and inputs. In concept, this study demonstrated that if screening of men at average risk, using a PSA threshold of 3.0 µg/L, was recommended to start at age 50 or 55 years, men with twice the risk of developing prostate cancer (disease onset) would derive similar benefit by screening starting at ages 45 or 48 years, respectively. If screening of men at average risk was recommended every 3 years, men with a 1.37 risk of progression of cancer would derive similar benefit by screening every 2 years. The model also demonstrated that men with an increased risk of developing cancer, screening strategies had a similar rank order as the general-risk population but the probability of preventing death due to prostate cancer and the frequency of overdiagnosis were both higher. In addition, for men with an increased risk of cancer progression, the frequency of overdiagnosis was lower and the probability of preventing prostate cancer death were higher than those of men in the general-risk population. Lowering the PSA threshold (i.e., 3.0 µg/L to 1.0 µg/L) did not improve outcomes in men at increased risk of onset or progression of prostate cancer.

The most recent study by Tsodikov and colleagues[44] used the three CISNET models, PSAPC, MISCAN-PRO, and SCANS, to examine the differential natural history of prostate cancer in AA men compared to the general U.S. population. The modelers first estimated the natural history in all races using SEER incidence and updated PSA screening data. Updated screening patterns indicate that fewer AA men received PSA testing (at least one PSA test) in all but the youngest ages compared to the general population through the 1990's. They then re-estimated the natural history in AA men by substituting PSA screening patterns for AA men, and re-estimated components of natural history (i.e., risk of disease onset and initial tumor features, risk of progression to metastasis and/or high-grade disease, and risk of clinical diagnosis). Re-estimating involved identifying values for each natural history component that allowed models to most closely match SEER prostate cancer incidence in AA men. They used a likelihood ratio test to evaluate whether re-estimating components of disease natural history significantly improved

the models' fits to the incidence data. They also reported the improvement in the goodness-of-fit by re-estimating components of natural history. The sequential estimation of natural history components found that allowing the risk of disease onset to be different for AA men provided an immediate improvement in the models' fits to incidence for this population. Allowing the risk of progression to distant stage to be different produced higher distant-stage but similar local-regional stage incidence projections. Allowing the risk of clinical diagnosis to differ provided modest improvements to the fit in some select cases.

Across the three models, AA men appeared to have more preclinical and progressive prostate cancer than the male general population. In the general population, the lifetime risk of developing preclinical disease was estimated at 25% to 29% versus 31% to 45% in AA men (24% to 54% higher).[44] The risk of clinical diagnosis was estimated at 25% to 66% higher in AA men than the general population. Among men who have had the disease onset, the risk of clinical diagnosis was estimated to be comparable for AAs (39% to 88%) and the general population (36% to 85%). This finding implied that the sojourn times from disease onset to diagnosis was very similar for AA men and the general population. However, among men with preclinical disease, the models estimated a 38% to 75% higher risk of metastasis before diagnosis in AA men compared to the general population, indicating a greater risk of progression (more aggressive disease) in AA men.

Based on these findings—in context of earlier modeling work suggesting that subpopulations at higher risk for onset of prostate cancer might benefit from earlier age to start screening and that subpopulations with more aggressive disease might benefit from shorter screening intervals—the authors suggested that initiating screening earlier and screening more frequently may result in greater benefit for AAs compared with the general population. The modeling publication did not present numeric estimates of the amount of increase in net benefits from earlier and more frequent screening of AA men compared with other men. The authors concluded that, if it is agreed that PSA screening is of value for the general population at age 55 years, the models suggest initiating screening 3 to 9 years earlier in AA men.

Summary

Thus far, modeling studies that address high-risk subpopulations are more conceptual than applied, demonstrating that within a general population, a population at higher risk for prostate cancer might benefit from earlier age to start screening if there is a greater risk of cancer onset in this population and/or a shorter interval of PSA screening if the cancer is more likely to progress faster, and assuming that treatments for more aggressive disease are equally effective at the same stage or grade. To date, modeling studies demonstrating the balance of benefits and harms for prostate cancer screening in AAs or those with a family history of (or genetic predisposition to) prostate cancer do not exist. Studies using robust estimates (with sensitivity analyses) for assumptions about disease incidence, stage distribution at screening, natural history (sojourn time, progression to metastatic disease), response to treatment (differential benefit or harm), and (dis)utilities around health states for important subpopulations at higher risk for prostate cancer morbidity or mortality would be helpful in the assessment of possible targeted screening recommendations.

Chapter 4. Discussion

Limitations

Our paper summarizes the most relevant modeling studies to current prostate cancer screening and treatment practice in the United States addressing four questions developed with guidance from the USPSTF. Our paper does not explicitly address the question of screening benefit from prostate cancer screening, as this has been primarily addressed by two large prostate cancer screening trials. Our paper does, however, address the net benefit (balance of benefits and harms) as this is arguably the more policy relevant question. Our analysis and summary of modeling studies is not a systematic review; however, we believe our overview represents and synthesizes both the highest quality and most applicable modeling studies to current U.S. practice. Given the nature of the questions asked in this paper, we focus on findings from modeling studies using necessarily complex models, and primarily those conducted by the three CISNET prostate cancer screening modeling groups. Our review does not address modeling studies on the value of risk prediction tools or adjunctive testing (to PSA testing) to aid in targeting or tailoring prostate cancer screening or subsequent diagnostic follow-up.

Conclusions

Given the potential implications of modeling studies for policy decisions regarding prostate cancer (i.e., whether or how to implement population-based or targeted screening), it is necessary to understand the findings of models in the context of the primary (trial) evidence and the limitations of the modeling exercises. Given the uncertainties and assumptions inherent to models, modeling studies are unlikely to give precise estimates of net benefit; however, these studies may offer a better understanding about critical factors to optimize screening strategies to mitigate harms, as well as estimates of upper and lower limits of the major benefits and harms of screening, with any given screening strategy. Modeling studies have demonstrated that screening strategies that optimize benefit (reduce prostate cancer mortality) increase potential harm (overdiagnosis). The estimated frequency of overdiagnosis varies widely but is substantial. Screening strategies that lower the age for stopping screening lengthen the interval between screenings; raise the PSA threshold for biopsy, and implement selective treatment (active surveillance) can all mitigate the potential harm of overdiagnosis and/or subsequent overtreatment. Strategies that mitigate harms lead to a decrement in benefit of reduction in prostate cancer mortality although the decrement can be small. Weighing the balance of benefits and harms of screening is subjective, and using integrated measures like QALYs is an evolving science. The evidence base for the measurement of (dis)utility of various health states related to prostate cancer screening, diagnosis, and treatment is limited, and estimates of (dis)utility vary widely. Lastly, the use of modeling exercises to understand the value of targeted screening strategies in groups at higher risk for prostate cancer and prostate cancer mortality is still conceptual, as studies evaluating the balance of benefits and harms in these subpopulations of men do not exist.

References

1. Moyer VA, US Preventive Services Task Force. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;157(2):120-34. PMID: 22801674. https://dx.doi.org/10.7326/0003-4819-157-2-201207170-00459
2. Etzioni R, Gulati R, Cooperberg MR, et al. Limitations of basing screening policies on screening trials: The US Preventive Services Task Force and Prostate Cancer Screening. Med Care. 2013;51(4):295-300. PMID: 23269114. https://dx.doi.org/10.1097/MLR.0b013e31827da979
3. Melnikow J, LeFevre M, Wilt TJ, et al. Counterpoint: Randomized trials provide the strongest evidence for clinical guidelines: The US Preventive Services Task Force and Prostate Cancer Screening. Med Care. 2013;51(4):301-3. PMID: 23481031. https://dx.doi.org/10.1097/MLR.0b013e31828a67d3
4. Armstrong EP. Prophylaxis of cervical cancer and related cervical disease: a review of the cost-effectiveness of vaccination against oncogenic HPV types. J Manage Care Pharm. 2010;16(3):217-30. PMID: 20331326. https://dx.doi.org/10.18553/jmcp.2010.16.3.217
5. Sox HC. Quality of life and guidelines for PSA screening. N Engl J Med. 2012;367(7):669-71. PMID: 22894580. https://dx.doi.org/10.1056/NEJMe1207165
6. Fenton JJ, Weyrich MS, Durbin S, et al. Prostate-Specific Antigen-Based Screening for Prostate Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force. Rockville, MD: Agency for Healthcare Research and Quality; 2017. PMID.
7. Boniol M, Boyle P, Autier P, et al. Critical role of prostate biopsy mortality in the number of years of life gained and lost within a prostate cancer screening programme. BJU Int. 2012;110(11):1648-52. PMID: 22984785. https://dx.doi.org/10.1111/j.1464-410X.2012.11513.x
8. Kobayashi T, Goto R, Ito K, et al. Prostate cancer screening strategies with re-screening interval determined by individual baseline prostate-specific antigen values are cost-effective. Eur J Surg Oncol. 2007;33(6):783-9. PMID: 17408910. https://dx.doi.org/10.1016/j.ejso.2007.02.015
9. Krahn MD, Mahoney JE, Eckman MH, et al. Screening for prostate cancer. A decision analytic view. JAMA. 1994;272(10):773-80. PMID: 7521400. https://dx.doi.org/10.1001/jama.1994.03520100035030
10. Lyth J, Andersson SO, Andren O, et al. A decision support model for cost-effectiveness of radical prostatectomy in localized prostate cancer. Scandinavian journal of urology and nephrology. 2012;46(1):19-25. PMID: 21905981. https://dx.doi.org/10.3109/00365599.2011.615759
11. Martin AJ, Lord SJ, Verry HE, et al. Risk assessment to guide prostate cancer screening decisions: a cost-effectiveness analysis. The Medical journal of Australia. 2013;198(10):546-50. PMID: 23725269. https://dx.doi.org/10.5694/mja12.11597
12. Nichol MB, Wu J, Huang J, et al. Cost-effectiveness of Prostate Health Index for prostate cancer detection. BJU Int. 2012;110(3):353-62. PMID: 22077934. https://dx.doi.org/10.1111/j.1464-410X.2011.10751.x

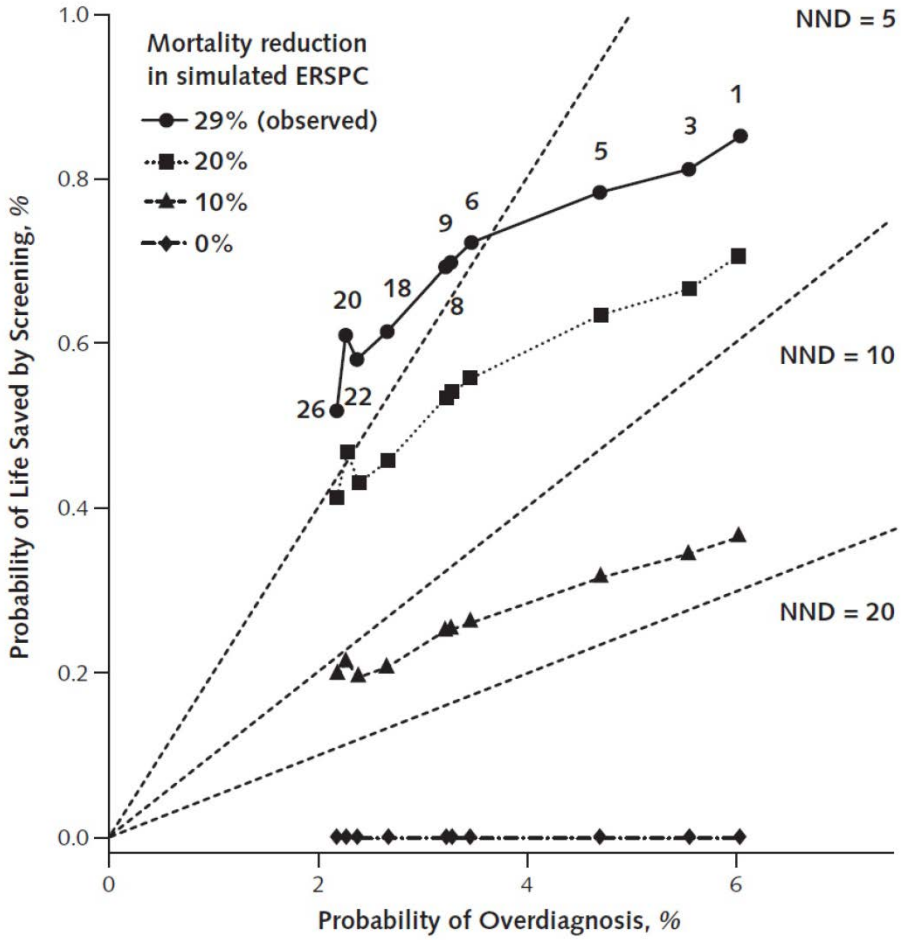
13. Ross KS, Carter HB, Pearson JD, et al. Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. JAMA. 2000;284(11):1399-405. PMID: 10989402. https://dx.doi.org/10.1001/jama.284.11.1399
14. Shteynshlyuger A, Andriole GL. Cost-effectiveness of prostate specific antigen screening in the United States: extrapolating from the European study of screening for prostate cancer. J Urol. 2011;185(3):828-32. PMID: 21239021. https://dx.doi.org/10.1016/j.juro.2010.10.079
15. Vickers A, Bennette C, Steineck G, et al. Individualized estimation of the benefit of radical prostatectomy from the Scandinavian Prostate Cancer Group randomized trial. Eur Urol. 2012;62(2):204-9. PMID: 22541389. https://dx.doi.org/10.1016/j.eururo.2012.04.024
16. Xia J, Trock BJ, Cooperberg MR, et al. Prostate cancer mortality following active surveillance versus immediate radical prostatectomy. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012;18(19):5471-8. PMID: 23008476. https://dx.doi.org/10.1158/1078-0432.CCR-12-1502
17. Pashayan N, Duffy SW, Pharoah P, et al. Mean sojourn time, overdiagnosis, and reduction in advanced stage prostate cancer due to screening with PSA: implications of sojourn time on screening. Br J Cancer. 2009;100(7):1198-204. PMID: 19293796. https://dx.doi.org/10.1038/sj.bjc.6604973
18. Wu GH, Auvinen A, Maattanen L, et al. Number of screens for overdetection as an indicator of absolute risk of overdiagnosis in prostate cancer screening. Int J Cancer. 2012;131(6):1367-75. PMID: 22052356. https://dx.doi.org/10.1002/ijc.27340
19. Wu GH, Auvinen A, Yen AM, et al. The impact of interscreening interval and age on prostate cancer screening with prostate-specific antigen. Eur Urol. 2012;61(5):1011-8. PMID: 22264679. https://dx.doi.org/10.1016/j.eururo.2012.01.008
20. Petitti D, Lin JS, Burda BU. Overdiagnosis in Prostate Cancer Screening Decision Models: A Contextual Review for the U.S. Preventive Services Task Force. Rockville, MD: Agency for Healthcare Research and Quality; 2017 May. PMID.
21. Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. Biometrics. 2008;64(1):10-9. PMID: 17501937. https://dx.doi.org/10.1111/j.1541-0420.2007.00825.x
22. Draisma G, De Koning HJ. MISCAN: estimating lead-time and over-detection by simulation. BJU Int. 2003;92 Suppl 2:106-11. PMID: 14983966. https://dx.doi.org/10.1111/j.1464-410X.2003.4409x.x
23. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst. 2003;95(12):868-78. PMID: 12813170.
24. McGregor M, Hanley JA, Boivin JF, et al. Screening for prostate cancer: estimating the magnitude of overdetection. CMAJ. 1998;159(11):1368-72. PMID: 9861205.
25. Davidov O, Zelen M. Overdiagnosis in early detection programs. Biostatistics. 2004;5(4):603-13. PMID: 15475422. https://dx.doi.org/10.1093/biostatistics/kxh012
26. Tsodikov A, Szabo A, Wegelin J. A population model of prostate cancer incidence. Statistics in medicine. 2006;25(16):2846-66. PMID: 16397859. https://dx.doi.org/10.1002/sim.2257
27. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst. 2009;101(6):374-83. PMID: 19276453. https://dx.doi.org/10.1093/jnci/djp001

28. Gulati R, Gore JL, Etzioni R. Comparative effectiveness of alternative prostate-specific antigen--based prostate cancer screening strategies: model estimates of potential benefits and harms. Ann Intern Med. 2013;158(3):145-53. PMID: 23381039. https://dx.doi.org/10.7326/0003-4819-158-3-201302050-00003
29. Roth JA, Gulati R, Gore JL, et al. Economic Analysis of Prostate-Specific Antigen Screening and Selective Treatment Strategies. JAMA Oncol. 2016;2(7):890-8. PMID: 27010943. https://dx.doi.org/10.1001/jamaoncol.2015.6275
30. de Carvalho TM, Heijnsdijk EA, de Koning HJ. Screening for prostate cancer in the US? Reduce the harms and keep the benefit. Int J Cancer. 2015;136(7):1600-7. PMID: 25123412. https://dx.doi.org/10.1002/ijc.29136
31. de Carvalho TM, Heijnsdijk EA, de Koning HJ. Estimating the individual benefit of immediate treatment or active surveillance for prostate cancer after screen-detection in older (65+) men. Int J Cancer. 2016;138(10):2522-8. PMID: 26695380. https://dx.doi.org/10.1002/ijc.29976
32. Hamdy FC, Donovan JL, Lane JA, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med. 2016. PMID: 27626136 https://dx.doi.org/10.1056/NEJMoa1606220
33. Stewart ST, Lenert L, Bhatnagar V, et al. Utilities for prostate cancer health states in men aged 60 and older. Med Care. 2005;43(4):347-55. PMID: 15778638.
34. Tsodikov A, Gulati R, Heijnsdijk EA, et al. Reconciling the effects of screening on prostate cancer mortality in the ERSPC and PLCO trials. Ann Intern Med. 2017. PMID: 28869989. https://dx.doi.org/10.7326/M16-2586
35. Wever EM, Hugosson J, Heijnsdijk EA, et al. To be screened or not to be screened? Modeling the consequences of PSA screening for the individual. Br J Cancer. 2012;107(5):778-84. PMID: 22805324. https://dx.doi.org/10.1038/bjc.2012.317
36. Heijnsdijk EA, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. N Engl J Med. 2012;367(7):595-605. PMID: 22894572. https://dx.doi.org/10.1056/NEJMoa1201637
37. Zhang J, Denton BT, Balasubramanian H, et al. Optimization of PSA screening policies: a comparison of the patient and societal perspectives. Med Decis Making. 2012;32(2):337-49. PMID: 21933990. https://dx.doi.org/10.1177/0272989x11416513
38. Underwood DJ, Zhang J, Denton BT, et al. Simulation optimization of PSA-threshold based prostate cancer screening policies. Health care management science. 2012;15(4):293-309. PMID: 22302420. https://dx.doi.org/10.1007/s10729-012-9195-x
39. Underwood DJ, Zhang J, Denton BT, et al. Note on "Simulation optimization of PSA-threshold based prostate cancer screening policies". Health care management science. 2013;16(4):377-8. PMID: 23640163. https://dx.doi.org/10.1007/s10729-013-9238-y
40. Heijnsdijk EA, de Carvalho TM, Auvinen A, et al. Cost-effectiveness of prostate cancer screening: a simulation study based on ERSPC data. J Natl Cancer Inst. 2015;107(1):366. PMID: 25505238. https://dx.doi.org/10.1093/jnci/dju366
41. Bremner KE, Chong CA, Tomlinson G, et al. A review and meta-analysis of prostate cancer utilities. Med Decis Making. 2007;27(3):288-98. PMID: 17502448. https://dx.doi.org/10.1177/0272989X07300604
42. Yen AM, Auvinen A, Schleutker J, et al. Prostate cancer screening using risk stratification based on a multi-state model of genetic variants. Prostate. 2015;75(8):825-35. PMID: 25683204. https://dx.doi.org/10.1002/pros.22964

43. Gulati R, Cheng HH, Lange PH, et al. Screening men at increased risk for prostate cancer diagnosis: model estimates of benefits and harms. Cancer Epidemiol Biomarkers Prev. 2016;26(2):222-7. PMID: 27742670. https://dx.doi.org/10.1158/1055-9965.EPI-16-0434
44. Tsodikov A, Gulati R, M. dCT, et al. Is prostate cancer different in black men? Answers from 3 natural history models. Cancer. 2017. PMID: 28436011. https://dx.doi.org/10.1002/cncr.30687
45. Laitinen VH, Wahlfors T, Saaristo L, et al. HOXB13 G84E mutation in Finland: population-based analysis of prostate, breast, and colorectal cancer risk. Cancer Epidemiol Biomarkers Prev. 2013;22(3):452-60. PMID: 23292082. https://dx.doi.org/10.1158/1055-9965.EPI-12-1000-T
46. Nurminen R, Wahlfors T, Tammela TL, et al. Identification of an aggressive prostate cancer predisposing variant at 11q13. Int J Cancer. 2011;129(3):599-606. PMID: 21064104. https://dx.doi.org/10.1002/ijc.25754
47. Bancroft EK, Page EC, Castro E, et al. Targeted prostate cancer screening in BRCA1 and BRCA2 mutation carriers: results from the initial screening round of the IMPACT study. Eur Urol. 2014;66(3):489-99. PMID: 24484606. https://dx.doi.org/10.1016/j.eururo.2014.01.003
48. Lansdorp-Vogelaar I, Gulati R, Mariotto AB, et al. Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. Ann Intern Med. 2014;161(2):104-12. PMID: 25023249. https://dx.doi.org/10.7326/m13-2867
49. Heijnsdijk EA, der Kinderen A, Wever EM, et al. Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer. Br J Cancer. 2009;101(11):1833-8. PMID: 19904272. https://dx.doi.org/10.1038/sj.bjc.6605422
50. Wever EM, Heijnsdijk EA, Draisma G, et al. Treatment of local-regional prostate cancer detected by PSA screening: benefits and harms according to prognostic factors. Br J Cancer. 2013;108(10):1971-7. PMID: 23674085. https://dx.doi.org/10.1038/bjc.2013.198
51. 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(13):981-90. PMID: 12096083.
52. Gulati R, Inoue L, Katcher J, et al. Calibrating disease progression models using population data: a critical precursor to policy development in cancer control. Biostatistics. 2010;11(4):707-19. PMID: 20530126. https://dx.doi.org/10.1093/biostatistics/kxq036
53. Gulati R, Inoue LY, Gore JL, et al. Individualized estimates of overdiagnosis in screen-detected prostate cancer. J Natl Cancer Inst. 2014;106(2):djt367. PMID: 24399850. https://dx.doi.org/10.1093/jnci/djt367
54. Pataky R, Gulati R, Etzioni R, et al. Is prostate cancer screening cost-effective? A microsimulation model of prostate-specific antigen-based screening for British Columbia, Canada. Int J Cancer. 2014;135(4):939-47. PMID: 24443367. https://dx.doi.org/10.1002/ijc.28732
55. Cooperberg MR, Ramakrishna NR, Duff SB, et al. Primary treatments for clinically localised prostate cancer: a comprehensive lifetime cost-utility analysis. BJU Int. 2013;111(3):437-50. PMID: 23279038. https://dx.doi.org/10.1111/j.1464-410X.2012.11597.x
56. Essink-Bot ML, de Koning HJ, Nijs HG, et al. Short-term effects of population-based screening for prostate cancer on health-related quality of life. J Natl Cancer Inst. 1998;90(12):925-31. PMID: 9637143.

57. de Haes JC, de Koning HJ, van Oortmarssen GJ, et al. The impact of a breast cancer screening programme on quality-adjusted life-years. Int J Cancer. 1991;49(4):538-44. PMID: 1917155.
58. Korfage IJ, de Koning HJ, Roobol M, et al. Prostate cancer diagnosis: the impact on patients' mental health. European journal of cancer (Oxford, England : 1990). 2006;42(2):165-70. PMID: 16326098. https://dx.doi.org/10.1016/j.ejca.2005.10.011
59. Konski A, Sherman E, Krahn M, et al. Economic analysis of a phase III clinical trial evaluating the addition of total androgen suppression to radiation versus radiation alone for locally advanced prostate cancer (Radiation Therapy Oncology Group protocol 86-10). International journal of radiation oncology, biology, physics. 2005;63(3):788-94. PMID: 16109464. https://dx.doi.org/10.1016/j.ijrobp.2005.03.010
60. Calvert NW, Morgan AB, Catto JW, et al. Effectiveness and cost-effectiveness of prognostic markers in prostate cancer. Br J Cancer. 2003;88(1):31-5. PMID: 12556955. https://dx.doi.org/10.1038/sj.bjc.6600630
61. Bennett CL, Matchar D, McCrory D, et al. Cost-effective models for flutamide for prostate carcinoma patients: are they helpful to policy makers? Cancer. 1996;77(9):1854-61. PMID: 8646685. https://dx.doi.org/10.1002/(SICI)1097-0142(19960501)77:9%3C1854::AID-CNCR15%3E3.0.CO;2-Z
62. Zeliadt SB, Etzioni RD, Penson DF, et al. Lifetime implications and cost-effectiveness of using finasteride to prevent prostate cancer. Am J Med. 2005;118(8):850-7. PMID: 16084177. https://dx.doi.org/10.1016/j.amjmed.2005.03.001
63. Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: progress and promise. J Clin Oncol. 2011;29(27):3669-76. PMID: 21825257. https://dx.doi.org/10.1200/JCO.2011.34.9738
64. Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. N Engl J Med. 2008;358(12):1250-61. PMID: 18354103. https://dx.doi.org/10.1056/NEJMoa074311
65. Konski A, Watkins-Bruner D, Brereton H, et al. Long-term hormone therapy and radiation is cost-effective for patients with locally advanced prostate carcinoma. Cancer. 2006;106(1):51-7. PMID: 16323171. https://dx.doi.org/10.1002/cncr.21575
66. Moeremans K, Caekelbergh K, Annemans L. Cost-effectiveness analysis of bicalutamide (Casodex) for adjuvant treatment of early prostate cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2004;7(4):472-81. PMID: 15449639. https://dx.doi.org/10.1111/j.1524-4733.2004.74010.x
67. Penson DF, Ramsey S, Veenstra D, et al. The cost-effectiveness of combined androgen blockade with bicalutamide and luteinizing hormone releasing hormone agonist in men with metastatic prostate cancer. J Urol. 2005;174(2):547-52; discussion 52. PMID: 16006889. https://dx.doi.org/10.1097/01.ju.0000165569.48372.4c
68. Ramsey S, Veenstra D, Clarke L, et al. Is combined androgen blockade with bicalutamide cost-effective compared with combined androgen blockade with flutamide? Urology. 2005;66(4):835-9. PMID: 16230148. https://dx.doi.org/10.1016/j.urology.2005.04.028

Figure 1. Tradeoff Between Probability of Life Saved and Overdiagnosis dor Selected Screening Strategies[28]



From Gulati R, Gore JL, Etzioni R. Comparative effectiveness of alternative prostate-specific antigen--based prostate cancer screening strategies: model estimates of potential benefits and harms. Ann Intern Med. 2013;158(3):145-53. Copyright © [2013] American College of Physicians. Reprinted with the permission of the American College of Physicians.

Each point represents the tradeoff for 10 of the 35 screening strategies examined in Gulati et al: the reference strategy (strategy 8), strategies that differ from the reference by a single screening variable (strategies 3, 5, 6, 9, 18, 20, and 26), and strategies based on recommendations by the National Comprehensive Cancer Network (strategy 1), the American Cancer Society (strategy 9), and Vickers and Lilja (strategy 22). The assumed effects of screening on prostate cancer survival correspond to mortality reductions of 29% (the reduction observed in the ERSPC trial after correction for noncompliance), 20%, 10%, and 0% projected in a simulated version of the ERSPC after 11 years of follow-up. Lines connect projections under the same mortality reduction.

The additional number needed to detect (NND) to prevent 1 prostate cancer death defined as the overdiagnoses divided by lives saved by screening. The NND corresponding to any point in the figure is obtained as the ratio of the probability of overdiagnosis to the probability of life saved. For reference, dashed lines radiating from the origin (representing no screening) illustrate fixed NND values of 5, 10, and 20. Different strategies will be preferred depending on relative weighting of the probabilities of life saved and overdiagnosis. Among the strategies considered, strategy 1 maximizes the probability of life saved and will be the preferred strategy if survival is the highest priority. Strategy 26 minimizes the probability of overdiagnosis and will be preferred if the morbidity associated with treatment is the greatest concern. For priorities between these extremes, the preferred strategy will be based on the most favorable balance between probabilities of life saved and overdiagnosis.

Table 1. Relevant Publications From Models of Prostate Cancer Screening

Model Name	Model Type	How Cancer Process Is Conceptualized	Population	Author, Year	Publication Title	Over-diagnosis	Strategies to Mitigate Harms	Utilities	Sub-populations
MISCAN-PRO (Erasmus)	Microsimulation	18 preclinical and clinical states as a semi-Markov process	United States	Draisma et al, 2009[27]	Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context.	X			
				Lansdorp-Vogelaar et al, 2014[48]	Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits.	X	X		
				de Carvalho et al, 2015[30]	Screening for prostate cancer in the US? Reduce the harms and keep the benefit.	X	X		
				Tsodikov, 2017[44]	Is prostate cancer different in black men? Answers from three natural history models				X
			Europe	Draisma et al, 2003[22, 23]	Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer.	X			
				Heijnsdijk et al, 2009[49]	Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer.	X			
				Heijnsdijk et al, 2012[36]	Quality-of-life effects of prostate-specific antigen screening.	X	X	X	
				Wever et al, 2012[35]	To be screened or not to be screened? Modeling the consequences of PSA screening for the individual.			X	
				Wever et al, 2013[50]	Treatment of local-regional prostate cancer detected by PSA screening: benefits and harms according to prognostic factors.	X	X		
				Heijnsdijk et al, 2015[40]	Cost-effectiveness of prostate cancer screening: a simulation study based on ERSPC data.	X	X	X	
PSAPC* (FHCRC)	Microsimulation	PSA growth by age superimposed on cancer growth proportional to PSA	United States	Etzioni et al, 2002[51]	Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends	X			
				Draisma et al, 2009[27]	Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context.	X			
				Gulati et al, 2010[52]	Calibrating disease progression models using population data: a critical precursor to policy development in cancer control.	X	X		

Table 1. Relevant Publications From Models of Prostate Cancer Screening

Model Name	Model Type	How Cancer Process Is Conceptualized	Population	Author, Year	Publication Title	Over-diagnosis	Strategies to Mitigate Harms	Utilities	Sub-populations
				Gulati et al, 2013[28]	Comparative effectiveness of alternative prostate-specific antigen--based prostate cancer screening strategies: model estimates of potential benefits and harms.	X	X		
				Gulati et al, 2014[53]	Individualized estimates of overdiagnosis in screen-detected prostate cancer.	X			
				Lansdorp-Vogelaar et al, 2014[48]	Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits.	X	X		
				Roth et al, 2016[29]	Economic Analysis of Prostate-Specific Antigen Screening and Selective Treatment Strategies.		X	X	
				Gulati et al, 2017[43]	Screening men at increased risk for cancer diagnosis: an integrate framework for policy development				X
				Tsodikov, 2017[44]	Is prostate cancer different in black men? Answers from three natural history models				X
			Canada	Pataky et al, 2014[54]	Is prostate cancer screening cost-effective? A microsimulation model of prostate-specific antigen-based screening for British Columbia, Canada.	X	X		
SCANS (Michigan)	Analytic	3-stage semi-Markov process based on chronic disease model	United States	Tsodikov et al, 2006[26]	A population model of prostate cancer incidence.	X			
				Draisma et al, 2009[27]	Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context.	X			
				Tsodikov, 2017[44]	Is prostate cancer different in black men? Answers from three natural history models				X
-- (NCSU)	Microsimulation	Markov process with 6 stages	United States	Underwood et al, 2012[38, 39]	Simulation optimization of PSA-threshold based prostate cancer screening policies.			X	
				Zhang et al, 2012[37]	Optimization of PSA screening policies: a comparison of the patient and societal perspectives.			X	

Table 1. Relevant Publications From Models of Prostate Cancer Screening

Model Name	Model Type	How Cancer Process Is Conceptualized	Population	Author, Year	Publication Title	Over-diagnosis	Strategies to Mitigate Harms	Utilities	Sub-populations
--	Macrosimulation /outcomes table	Incident prostate cancer cases progress to death due to prostate cancer based on stage-specific survival or death occurs from competing causes	Canada	McGregor et al, 1998[24]	Screening for prostate cancer: estimating the magnitude of overdetection.	X			
--	Analytic	3-stage process based on chronic disease model	United States	Davidov et al, 2004[25]	Overdiagnosis in early detection programs.	X			
--	Analytic	Does not conceptualize natural history of cancer; rather, focuses on the effects of screening on incidence	United States	Telesca et al, 2008[21]	Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends.	X			
--	Microsimulation	6-state Markov process	Europe (Finland)	Yen et al, 2015[42]	Prostate cancer screening using risk stratification based on a multi-state model of genetic variants.				X

*The FHCRC model, previously called Prostate Cancer SIMulation (PCSIM), was renamed PSAPC in 2009, after a new parametrization was incorporated to facilitate empirical estimation of key relationships, including relationships between PSA growth and cancer progression.

Abbreviations: FHCRC = Fred Hutchinson Cancer Research Center; N/A = not applicable; MISCAN-PRO = MIcrosimulaton SCreening ANalysis Prostate Cancer model; NCSU = North Carolina State University; PSA = prostate-specific antigen; PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression; UK = United Kingdom; US = United States

Table 2. Overview of CISNET Prostate Cancer Screening Models

Model Institution Type	Population Main Data Sources	Inputs and Assumptions Natural History	 Screening	 Treatment	Outcomes
MISCAN- PRO Erasmus Microsimulation	U.S. and Europe (Rotterdam trial, Sweden trial, Netherlands Dutch) Registries-SEER, Dutch and Swedish cancer registries Trials (ERSPC, PLCO, SPCG-4) Observational studies-CaPSURE Survey (NHIS) Other-Country-specific life tables including US life tables (1903-1959)	Cancer growth rates from ERSPC-Rotterdam or PLCO Natural and clinical history estimated via calibration to cancer registries or SEER data (1975-2000) 18 preclinical detectable states derived from combinations of clinical T-stage (T1-3), Gleason grade (well, moderately, and poorly differentiated), and metastatic stage (local-regional and distant) Assumes all cancers begin in localized stage and progress to metastasis All cancers begin in low grade and can progress over time Does not allow for cancer recurrence Death from prostate cancer and death from other causes are independent Lifetime risk of prostate cancer is the same for all men in the same birth cohort	Screening dissemination parameters from ERSPC-Rotterdam or NHIS, SEER-Medicare (2000) Cure rate from screening (and subsequent treatment) from ERSPC PSA screening and subsequent biopsy modeled as one single test, such that test sensitivity combines the probability of a positive PSA test, receipt of biopsy, and sensitivity of the biopsy to detect latent cancer Mechanism for survival benefit: the effect of lead time does not drive the estimate of survival benefit; a part of the screen-detected men is cured from cancer and that for the other part does not alter the life history (mortality benefit calibrated to ERSPC trial)	Treatment dissemination data from ERSPC-Rotterdam or SEER and CaPSURE Treatment benefit from SPCG-4 and observational studies Treatment benefit affected by temporal trend in calendar year in studies of mortality trends but not comparative effectiveness of candidate screening strategies Benefit depends on treatment modality (includes conservative management, radical prostatectomy, and radiation therapy +/- androgen deprivation therapy)	Cancer incidence Survival (life-years), QALY, mortality Harms (false positive results, unnecessary biopsies, overdiagnoses), cost

Table 2. Overview of CISNET Prostate Cancer Screening Models

Model Institution Type	Population Main Data Sources	Inputs and Assumptions			Outcomes
		Natural History	Screening	Treatment	
PSAPC* FHCRC Microsimulation	U.S. Registries-SEER, SEER-Medicare Trials- ERSPC, PCPT, PLCO, SPCG-4 Survey- NHIS Other- US life tables (1903-1959)	Cancer and PSA growth rates from PCPT and PLCO Natural and clinical history estimated via calibration to SEER data (1975-2000) Models longitudinal PSA growth and 3 natural history states (healthy, preclinical, clinical) or 9 states if 2 stage and 2 grade subcategories in preclinical and clinical states accounted for) which are dependent on age and PSA growth Assumes all cancers begin in localized stage and progress to metastasis Cancers can be low or high grade at onset but cannot progress over time† Does not allow for cancer recurrence Death from prostate cancer and death from other causes are independent PSA growth is log-linear in age, change point occurs at onset, growth rates are heterogeneous across individuals, differs with high- and low-grade disease	Screening dissemination parameters from NHIS, SEER-Medicare (2000) Biopsy compliance data from PLCO (depends on age and PSA at diagnosis) Generates PSA level per individual at each screen, PSA >4 µg/L at screen referred to biopsy Biopsy sensitivity increases with dissemination of extended biopsy schemes over time Mechanism for survival benefit from early detection (in part or in whole) from stage shift, hence the effect of screening on survival benefit depends on lead time (stage shift consistent with mortality reduction in ERSPC)	Treatment dissemination data from SEER (1975-2005 or 2010) Treatment benefit from SPCG-4 and observational studies Treatment benefit affected by temporal trend in calendar year and age in studies of mortality trends but not comparative effectiveness of candidate screening strategies Benefit depends on treatment modality (includes conservative management, radical prostatectomy, radiation therapy, and androgen deprivation therapy)	Cancer incidence Screening test performance Survival (life-years), QALY, mortality Harms (false positive results, false negative results, unnecessary biopsies, overdiagnoses), cost

Table 2. Overview of CISNET Prostate Cancer Screening Models

Model Institution Type	Population Main Data Sources	Inputs and Assumptions Natural History	Screening	Treatment	Outcomes
SCANS University of Michigan Analytic mathematical model	U.S. Registries-SEER Trials- ERSPC, PLCO Survey- NHIS Other- US life tables (1903-1959)	Natural and clinical history from SEER (1975-2000) 3 natural history states (healthy, preclinical, clinical) Does not specify stage or grade of tumor at onset Allows for cancer recurrence‡ Death from prostate cancer and death from other causes are independent	Screening dissemination parameters NHIS, SEER-Medicare (2000 or 2005) PSA screening and subsequent biopsy modeled as one single test such that test sensitivity combines the probability of a positive PSA test, receipt of biopsy, and sensitivity of the biopsy to detect latent cancer Sensitivity of screening test in an increasing function of time since tumor onset Mechanism for survival benefit from early detection (in part or in whole) from stage shift, hence, the effect of screening on survival benefit depends on lead time	Treatment dissemination data from SEER (1975-2005 or 2010) Treatment benefit from SPCG-4 and observational studies Treatment benefit affected by temporal trend in calendar year, age and birth cohort in studies of mortality trends but not comparative effectiveness of candidate screening strategies Benefit depends on treatment modality (NR)	Cancer incidence Survival (life-years), mortality Harms (overdiagnoses)

*The FHCRC model, previously called Prostate Cancer SIMulation (PCSIM), was renamed PSAPC in 2009, after a new parametrization was incorporated to facilitate empirical estimation of key relationships, including relationships between PSA growth and cancer progression.
†Low-moderate (Gleason 2-7) versus high (Gleason 8-10), recent model development to distinguish Gleason 2-6 versus 7
‡In studies of population incidence and mortality trends, does not explicitly include recurrence

Abbreviations: CaPSURE = Cancer of the Prostate Specific Urologic Research Endeavor; ERSPC = European Randomized Study of Screening for Prostate Cancer; FHCRC = Fred Hutchinson Cancer Research Center; NHIS = National Health Interview Survey; MISCAN-PRO = MIcrosimulaton SCreening ANalysis PROstate Cancer model; NR = not reported; PCPT = Prostate Cancer Prevention Trial; PLCO = Prostate, Lung, Colorectal and Ovarian Cancer screening trial; PSA = prostate-specific antigen; PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression; QALY = quality-adjusted life-year; SEER = Surveillance, Epidemiology, and End Results; SPCG = Scandinavian Prostate Cancer Group; US = United States

Table 3. Estimated Lifetime Benefits and Harms of Screening for Prostate Cancer for Selected PSA Screening Strategies, for the U.S. Population Using the PSAPC Model (Strategies Ordered According to Probability of Prostate Cancer Deaths)[28]

Screening Strategy	Age to Start/Stop Screening, years	Screening Interval	PSA Threshold for Biopsy, µg/L	Probability of Overdiagnosis (as a Proportion of Men Screened), %	Probability of Prostate Cancer Death, %	Probability of Life Saved, %	Number Needed to Detect to Prevent One Prostate Cancer Death**
1 (NCCN)	40-74	Complex†	Complex‡	6.0	2.02	0.85	7.08
3	50-74	Annual	Complex*	5.5	2.05	0.81	6.84
5	50-74	Annual	>2.5	4.7	2.08	0.78	6.01
6	40-74	Annual	>4.0	3.5	2.13	0.72	4.79
8 (reference)	50-74	Annual	>4.0	3.3	2.15	0.70	4.70
9 (ACS)	50-74	Complex§	>4.0	3.3	2.15	0.70	4.70
11	50-74	Biennial	>2.5	3.8	2.16	0.69	5.51
18	50-74	Biennial	>4.0	2.7	2.23	0.61	4.34
20	50-74	Annual	Complex¶	2.3	2.23	0.61	3.71
21	50-69	Annual	>2.5	2.9	2.24	0.61	4.75
22 (Vickers & Lilja)	45-74	Complex‖	>4.0	2.4	2.27	0.58	4.09
26	50-69	Annual	>4.0	1.8	2.32	0.51	3.58
30	50-69	Biennial	>2.5	2.0	2.35	0.49	4.12
35	50-69	Biennial	>4.0	1.3	2.43	0.41	3.11

*PSA > 4.0 µg/L or PSA velocity >0.35 µg/L per year
†Annual (quinquennial if age <50 and PSA <1 µg/L)
‡PSA >2.5 µg/L or PSA velocity >0.35 µg/L per year
§Annual (biennial if PSA <2.5 µg/L)
‖Bienniel (quinquennial if PSA < median for age)
¶>95th percentile for age
**Probability of overdiagnosis divided by life saved

Abbreviations: PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression; ACS = American Cancer Society; L = liter(s); NCCN = National Comprehensive Cancer Network; PSA = prostate-specific antigen; µg = microgram(s)

Table 4. Estimated (per Person) Life-Years and Quality-Adjusted Life-Years of Screening for Prostate Cancer for Selected Contemporary and Selective Treatment PSA Screening Strategies, for the U.S. Population Using the PSAPC Model (Strategies Ordered According to Quality-Adjusted Life-Years)[29]

Screening Strategy	Age to Start/Stop Screening, years	Screening Interval	PSA Threshold for Biopsy, µg/L	Contemporary Treatment Scenario			Selective Treatment Scenario		
				LY	QALY	Incremental QALY*	LY	QALY	Incremental QALY*
No screening	-	-	-	36.302	21.504	--	36.302	21.504	--
4	45-69	Complex§	10.0	36.347	21.508	0.004	NA	NA	NA
18	55-69	Quinquennial	10.0	36.329	21.508	0.004	NA	NA	NA
12	50-74	Quinquennial	10.0	36.338	21.507	0.003	NA	NA	NA
11	50-74	Complex‖	10.0	36.348	21.507	0.003	NA	NA	NA
9	50-74	Annual	10.0	36.357	21.507	0.003	NA	NA	NA
17	55-69	Biennial	10.0	36.338	21.507	0.003	NA	NA	NA
3	45-69	Annual	10.0	36.345	21.507	0.003	NA	NA	NA
16	55-69	Annual	10.0	36.343	21.506	0.002	NA	NA	NA
15	55-69	Quinquennial	3.0	36.343	21.502	-0.002	36.338	21.508	0.004
8	50-74	Quinquennial	4.0	36.348	21.502	-0.002	36.343	21.508	0.004
10	50-74	Annual	Based on age†	36.361	21.502	-0.002	36.355	21.509	0.005
1	45-69	Annual	4.0	36.361	21.499	-0.005	36.354	21.509	0.005
7	50-74	Complex‖	4.0	36.359	21.499	-0.005	36.352	21.508	0.004
6	50-74	Annual	Based on age‡	36.363	21.498	-0.006	36.357	21.508	0.004
13	55-69	Annual	4.0	36.355	21.498	-0.006	36.350	21.508	0.004
14	55-69	Biennial	3.0	36.353	21.498	-0.006	36.349	21.508	0.004
5	50-74	Annual	4.0	36.366	21.494	-0.01	36.360	21.507	0.003
2	45-69	Complex§	3.0	36.360	21.494	-0.01	36.353	21.506	0.002

*Compared with no screening; a negative number implies a decrement in QALYs compared with no screening
†PSA thresholds for biopsy referral are 4.5, 5.5, and 8.5 µg/L for ages 50-59 years, 60-69 years, and 70-74 years, respectively
‡PSA thresholds for biopsy referral are 3.5, 4.5, and 6.5 µg/L for ages 50-59 years, 60-69 years, and 70-74 years, respectively
§Annual if PSA >3.0 µg/L otherwise biennial
‖Biennial if PSA >1.0 otherwise quinquennial

Abbreviations: PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression; LY = life year(s); L = liter(s); NA = not applicable; PSA = prostate specific antigen; QALY = quality-adjusted life year(s); µg = microgram(s)

Table 5. Estimated (per 1,000 Men) Life-Years and Quality-Adjusted Life Years of Screening for Prostate Cancer for Selected PSA Screening Strategies, for the European Population Using the MISCAN-PRO Model[36]

Screening Strategy	Age to Start/Stop Screening, years	Screening Interval	PSA Threshold for Biopsy, µg/L*	LYs Gained per 1,000 Men†	QALYs Gained per 1,000 Men†
1 (base model)	55-69	Annual	3.0	73	56
2	55-74	Annual	3.0	85	56
3	55-69	Annual	3.0	52	41
4	55	Once	3.0	12	12
5	60	Once	3.0	22	19
6	65	Once	3.0	25	17

*PSA threshold of 3.0 µg/L (range, 2.5 to 4.0 µg/L)
†Compared to no screening

Abbreviations: MISCAN-PRO = Microsimulaton Screening Analysis Prostate Cancer model; LY = life year(s); L = liter(s); NA = not applicable; PSA = prostate specific antigen; QALY = quality-adjusted life year(s); µg = microgram(s)

Appendix Table 1. Estimated Lifetime Benefits and Harms of Screening for Prostate Cancer for Selected PSA Screening Strategies, for the U.S. Population Using the MISCAN-PRO Model (Strategies Ordered According to Probability of Prostate Cancer Deaths)[30]

Maximum Prostate Cancer Mortality Increase Threshold	Age to Start/Stop Screening, years	Screening Interval	PSA Threshold for Biopsy, µg/L	Probability of Overdiagnosis (as a Proportion of Men Screened), %	Probability of Prostate Cancer Death (as a Proportion of Men Screened), %
0% (base case)	50-74	Annual	3.0	3.80	2.38
1%	50-74	Complex*	3.0	3.58	2.39
2%	50-72	Annual	3.0	3.12	2.41
3%	50-72	Biennial	3.0	2.88	2.45
4%	50-70	Annual	3.0	2.51	2.46
5%	50-70	Annual	Based on age†	2.15	2.49
7%	50-70	Quinquennial	3.0	1.99	2.53

*Annual (if age greater than 65 years and PSA < 1 µg/L, quinquennial)
†3.0 µg/L (if age greater than 66 years, PSA 4.0 µg/L)

Abbreviations: MISCAN-PRO = Microsimulaton Screening Analysis Prostate Cancer model; L = liter(s); PCM = prostate cancer mortality; PSA = prostate specific antigen; µg = microgram(s)

Appendix Table 2. Values of Health State Utilities Used for PSAPC Model Inputs by Roth et al, 2016[29]

Health State Variable	Point Estimate	Range
Healthy	1.00	0.90-1.00
Utility Decrement		
Symptomatic	0.11	0.05-0.17
Surveillance	0.08	0.02-0.14
Short-term treatment	0.25	0.19-0.31
Long-term treatment	0.08	0.02-0.14
Distant stage	0.25	0.22-0.28
End of life	0.67	0.57-0.77

Notes: All utility decrements with the exception of "healthy state" utility which was assumed are derived from Stewart and colleagues, 2005[33] The utility decrements are anchored to the 'healthy state' utility. The utility decrements from Stewart and colleagues are set up relative to the original 0.92 utility for the 'healthy' state.[55] So, the utility decrement for 'surveillance' is 0.92 minus the mean utility from Table 2 in Stewart and colleagues for cancer with 20% chance of spread (0.84), therefore 0.92-0.84=0.08. The decrement for 'symptomatic' is calculated similarly using the mean utility for cancer with 40% chance of spread (0.81), therefore 0.92-0.81=0.11. The decrement for 'short term treatment' is calculated using the prostatectomy utility (0.67), therefore 0.92-0.67=0.25. The decrement for 'long term' is assumed to be the same as 'surveillance' (0.08) given the assumption that 'long term' state only involves surveillance. 'Distant stage' is the cancer spread (asymptomatic) utility (0.67), therefore 0.92-0.67=0.25. 'End of life' is the metastatic utility (0.25), therefore 0.92-0.25=0.67. (Email communication with Joshua A. Roth, Ph.D. (November 8, 2016) to discuss utilities outcomes in Roth JA, *JAMA Oncol.* 2016).

Appendix Table 3. Values of Health State Utilities Used for MISCAN-PRO Model Inputs by Heijnsdijk et al, 2012, 2016 and 2015[35, 36, 40]

Health State Variable	Base-case	Favorable	Unfavorable
Screening attendance[56, 57]	0.99	1.00	0.99
Biopsy (diagnostic phase)[57]	0.90	0.94	0.87
Cancer diagnosis[58]	0.80	0.85	0.75
Radiation therapy: at 2 months after procedure[33]	0.73	0.91	0.71
Radiation therapy: at > 2 months to 1 year after procedure[59]	0.78	0.88	0.61
Radical prostatectomy: at 2 months after procedure[33]	0.67	0.90	0.56
Radical prostatectomy: at > 2 months to 1 year after procedure[60]	0.77	0.91	0.70
Active surveillance[61-63]	0.97	1.00	0.85
Post-recovery period[33, 64]	0.95	1.00	0.93
Palliative therapy[65-68]	0.60	0.24	0.86
Terminal illness[65, 67, 68]	0.40	0.24	0.40

Sources for utilities are provided in the reference list.