Contextual Review

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

AHRQ Publication No. 17-05229-EF-3
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, Eric (Rocky) Feuer, PhD, Joshua J. Fenton, MD, MPH, and Joy Melnikow, MD, MPH, for providing expert review of the draft report; Smyth Lai, MLS, and Elizabeth Hess, MA, ELS(D) at the Kaiser Permanente Center for Health Research.

Suggested Citation

Petitti D, Lin JS, Burda BU. Overdiagnosis in Prostate Cancer Screening Decision Models: A Contextual Review for the U.S. Preventive Services Task Force. AHRQ Publication No. 17-05229-EF-3. 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. Because empirical data to estimate overdiagnosis is quite limited for prostate cancer screening, modeling data can be helpful in bounding estimates of overdiagnosis and exploring how altering or targeting screening strategies could mitigate the frequency of overdiagnosis. 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 era of screening with prostate-specific antigen was sizeable. Estimates of overdiagnosis are dependent on many factors which vary across different populations (e.g., natural history of cancer, life expectancy of the population, differences in clinical practice, and prior history of screening). Modeling studies suggest that raising the PSA threshold, lengthening the interval of screening, and lowering the age to stop screening would all reduce the frequency of overdiagnosis. 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.

Chapter 1. Introduction

The problem of overdiagnosis has received an increasing amount of attention in the field of cancer screening. It is a particularly large issue for prostate cancer because prostate cancer can have a long period when it is detectable but asymptomatic. As early as the 1980's, before the use of prostate-specific antigen (PSA) for early detection of prostate cancer, overdiagnosis was recognized as an important issue for prostate cancer screening that used digital rectal examinations.[1]

Individuals with cancer that is overdiagnosed do not benefit from having their cancer detected by screening but they suffer from the harms of evaluations done to establish that cancer exists and the harms of treatment for the cancer. The harms of treating prostate cancer can be both serious and common.[2, 3] Even if individuals with screen-detected prostate cancer do not undergo treatment, they may suffer from anxiety and diminished well-being because of the cancer diagnosis and they may be burdened by the testing and interventions used to monitor the cancer. Uncertainties regarding the benefit of PSA screening in reducing prostate cancer mortality in randomized trials[2] and evidence that the incidence of prostate cancer increased dramatically with the adoption of PSA screening,[4-6] have moved overdiagnosis to the forefront of discussions about whether and how to screen for prostate cancer.

Our aim for this paper is to describe the evolution of findings about overdiagnosis in screening for prostate cancer and to glean information from the model-based literature that might be useful in thinking about the design of screening programs that would mitigate harms due to overdiagnosis and/or in understanding the uncertainties about overdiagnosis.

Historical Overview of Modeling Studies to Evaluate Screening Programs and Overdiagnosis

Modeling studies have been used to evaluate cancer screening programs for many decades. The 1969 the seminal work of Zelen and Feinleib,[7] which underpins the current modeling work on prostate cancer screening described a relatively simple model to estimate the time by which a screening (or diagnostic) test advances detection of a chronic disease—the lead time. They used their model and information from the Health Insurance Plan (HIP) randomized trial of breast cancer screening to estimate that the mean lead time for breast cancers detected at the first screening examination in the HIP trial was 2.36 years, as opposed to 20 months, based on the well-known epidemiologic relationship between the prevalence, incidence, and duration of breast cancer.[7] The model-derived estimate of lead time was then used to improve on the evaluation of the effect of screening on mortality due to breast cancer in the HIP trial. Walter and Day[8] extended the work of Zelen and Feinleib[7] by showing that the model-based estimate of the mean lead time is sensitive to assumptions about its distribution (Zelen and Feinleib[7] had assumed an exponential distribution). Walter and Day[8] used three distributions to fit the HIP data, the exponential distribution, the log normal distribution, and an empiric step function; they found that an exponential distribution for lead time provided the best fit to the HIP trial data. All

subsequent model-based evaluations of overdiagnosis derive from the work of these early modelers.

Since then, models used to evaluate cancer screening have adopted increasingly complex conceptual representations of the cancer process. Simulation has moved to the forefront of the modeling field as a tool for estimation of model outcomes. These models have been used to assess increasingly complex questions about prostate cancer screening, including how prostate cancer screening programs have affected, or could affect, the frequency of overdiagnosis. Over time, the same models and their assumptions have been changed in efforts to improve their ability to address new questions, incorporate better empiric information, increase their flexibility, and make them more representative of contemporary experience.

Key Factors That Influence Estimation of Overdiagnosis

The literature about overdiagnosis presents interpretive challenges because there are many factors that influence overdiagnosis in cancer screening programs and many factors that affect estimates of overdiagnosis derived from both modeling and other types of studies.[9] These factors include the definition used to conceptualize overdiagnosis, the metric used to measure the frequency of overdiagnosis, the context in which information about overdiagnosis is collected or applied, and the types of study designs and estimation approaches used to assess overdiagnosis. In seeking information about the frequency of overdiagnosis in a policy context, the multiplicity of factors that may influence overdiagnosis estimates is not always appreciated. Etzioni and colleagues[9] discussed factors that affect estimates of overdiagnosis for breast and prostate cancer screening in considerable detail. Here we discuss pertinent key factors when considering modeling studies of prostate cancer overdiagnosis.

Definition of Overdiagnosis

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. Alternative definitions of overdiagnosis include: 1) screen-detected cancer that would not have caused death due to prostate cancer before death from another cause and 2) cancer that would never progress no matter how long the patient lived (and within the limits of the maximum human lifespan) or would regress. A clear statement of the definition of overdiagnosis being used is sometimes absent from publications, which muddies the literature and presents opportunities to underplay or overplay overdiagnosis.

Metric Used to Measure Overdiagnosis

The "frequency" of overdiagnosis is generally presented as a proportion. The use of proportions presents many options for the choice of numerator and denominator.[9] For prostate cancer studies, the most common numerator used to measure the frequency of overdiagnosis is the number of prostate cancers detected by screening that would not have been clinically detected before the death of the individual from other causes. The most common denominators used are: 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. A measure of the frequency of overdiagnosis using the third denominator estimates the lifetime risk of overdiagnosis; it is often referred to as the probability of overdiagnosis. Many publications use more than one metric to report overdiagnosis.

For any given estimate of the number of overdiagnosed prostate cancers (i.e., the numerator), the frequency of overdiagnosis based on any one of these denominator definitions is different but predictably related. A measure of the frequency of overdiagnosis using the first denominator (i.e., estimated number of overdiagnosed cancers divided by the number of screen-detected cancers) will be higher than a measure that uses the second denominator (i.e., estimated number of overdiagnosed cancers divided by the number total of cancers detected) because screen-detected prostate cancers are a subset of all detected prostate cancers. Both of these measures will be higher than a measure of overdiagnosis based on the third denominator (i.e., estimated number of overdiagnosed cancers divided by the number of men screened) because only a small percentage of men screened will have prostate cancer. With the same numerator, estimates of the frequency of prostate cancer overdiagnosis based on different denominator definitions can differ by an order of magnitude; therefore, it is essential to pay careful attention to the actual metric being used. However, in the modeling literature, the metrics used to describe the frequency of overdiagnosis are inconsistent and may not be easily discernible in publications.[9-11]

Context of the Populations

The natural history of prostate cancer—the likelihood of cancer arising and the speed with which prostate cancer progresses in the absence of treatment—differs between populations, evidenced by differences in prostate cancer incidence and mortality, stage at presentation, and in the clinical and pathologic characteristics of prostate cancer between populations. For a prostate cancer screening test with a given level of sensitivity, the frequency of overdiagnosis depends on the natural history, which may vary across populations (e.g., Japan versus United States) or in subgroups within a population (e.g., by race/ethnicity).

For screen-detected cancers with the same lead time and prognosis, the frequency of overdiagnosis will be greater in a population with a shorter rather than longer life expectancy. The life expectancy at age 50 years for men currently varies by country (from 69 to 70 years to 82 to 83 years.[12] Additionally today's life expectancy is not the same as it was in the past. For example, in the United States, the estimated life expectancy at age 50 years for men was to age 73.1 years in 1955 but to age 79.2 years in 2006, an increase of 6.1 years. And to follow, current life-expectancy is not what it will be in the future.[12] Model-based estimates of overdiagnosis involve projecting population history forward.[13] Co-morbidity also affects life expectancy, and the frequency of overdiagnosis will vary according to co-morbidity. Models of overdiagnosis must attend carefully to the use of the appropriate data for estimating life expectancy and should be framed in relation to uncertainties about future life expectancy. The frequency of overdiagnosis will also depend on the prior history of screening. For example, among men age 65 years in 2016 the frequency of overdiagnosis will be different (and higher) for those who undergo a first screening for prostate cancer than for those who had been screened biennially

starting at age 50 years. Within a population, the frequency of overdiagnosis depends on the history of uptake of screening in the past and therefore the frequency changes over time. Other aspects of the clinical approach to screening and follow-up (e.g., the PSA threshold for biopsy referral, the proportion of men with a positive PSA tests who choose to have a biopsy, and the number of cores), will also affect the frequency of overdiagnosis.

Study Design and Estimation Approaches

A final factor affecting estimates of overdiagnosis is study design and estimation approaches. Estimates of overdiagnosis have been reported based on pathologic (autopsy) studies (which we do not discuss further); follow-up of randomized trials of screening; population data; and statistical, analytic, microsimulation, and decision analysis modeling studies.[14] Estimation approaches are the excess incidence approach and the lead time approach. Studies of overdiagnosis based on randomized trials and population studies use the excess incidence approach and sometimes incorporate modeling of screening. The modeling studies we discuss use the lead time approach; lead time is derived as an intermediate estimate from a model.

The excess incidence approach is conceptually simple. The number of prostate cancers observed in a screened population is compared with the number expected without screening; the difference is the number of overdiagnosed cancers. As we described earlier, the frequency of overdiagnosis can be expressed using one of three denominators: screen-detected prostate cancers, total prostate cancers, total number of men screened. Using the excess incidence approach, data about the number of observed and expected cancers in the screened population can be derived from randomized trials or population studies. When data from randomized trials are used, the difference between the number of cases of prostate cancer in the screened population and the number of cases observed in the control population yields the estimated number of excess, overdiagnosed cancers. When data from populations studies are used, the number of incident cases of prostate cancer observed over a specific period time after introduction of screening is compared with the number expected had there been no screening.

The conceptual simplicity of the excess incidence approach to assess the frequency of overdiagnosis belies the many difficulties when used in practice. To obtain an unbiased estimate of overdiagnosis when using clinical trial data and the excess incidence approach, it is critical to have a sufficient amount of follow-up time of trial participants. This is minimally the time until the screening rate stabilizes plus the maximum lead time.[14] To obtain a perfect estimate of overdiagnosis from randomized trials, it would be necessary to follow the subjects in the trial until all of them had died and to prohibit screening in the control population until then. When randomized trials are the source of data on excess incidence, the amount of screening within the trial for those assigned to the control group, the duration of screening in the trial, the amount of continued screening after trial end for those assigned to be screened within the trial, the uptake of screening in the control group after the trial ends, and the duration of follow-up will all affect estimates of overdiagnosis and can lead to bias in these estimates.[14] Information is also required on test sensitivity and the duration of the preclinical phase. Using population data to estimate overdiagnosis involves projecting background incidence trends forward in time, often over a long period, and requires accurate data on patterns of screening in the population, again over a long period of time.[9, 14, 15] Estimating the background trend for prostate cancer in the absence of

screening is complicated because the trend is affected by changes in clinical practice over time (e.g., medical treatment vs. transurethral prostatectomy for benign prostatic hypertrophy, changes in the number of cores in a biopsy).

The lead time approach to estimate overdiagnosis is also conceptually simple. This approach is illustrated best with a hypothetical example involving an extremely elderly man undergoing PSA screening. Imagine that a man age 115 years has his first PSA screening test, it is positive, and he is found to have prostate cancer. If the screening test advances the diagnosis of cancer by 6 years (i.e., the lead time—the time between detection by screening and when the cancer would have been detected clinically—is 6 years), the chances that this man's cancer is overdiagnosed is practically 100% because the chances that this man will live to 120 years is close to zero. A model simply quantifies the probability that death precedes diagnosis at the end of his lead time. Using the lead time approach, models generate estimates of lead time and age at screen detection, couple this information with empiric data on life expectancy, and generate estimates of the chances of overdiagnosis for a hypothetical population based on the conceptual reasoning in our example.

Like the excess incidence approach, the lead time approach is bedeviled by numerous complexities in execution. To estimate the lead time and its distribution, the model must disentangle the time of onset, the time of clinical diagnosis, and test sensitivity based on the natural history, which is unobserved. Estimating time of onset and clinical diagnosis requires context-specific information on incidence and screening patterns. Data on incidence and screening patterns may not be perfect.

Chapter 2. Methods

Analysis of Modeling Studies That Addressed Overdiagnosis in Prostate Cancer Screening

Consumers of information about prostate cancer overdiagnosis often ask how frequent overdiagnosis is for prostate cancer screening. There is no single answer to this question.[9] In approaching the literature about overdiagnosis from modeling studies, we did not attempt to find the answer to this question. While we have tried to be as complete as possible in identifying modeling studies that reported information about overdiagnosis, our synthesis is not a systematic review. We do not formally critically appraise models but instead we identified and described the most important assumptions related to overdiagnosis. Our overarching aim was to explicate the evolution of findings about overdiagnosis from modeling studies and to glean information from the literature that might be useful in thinking about the design of programs that would mitigate harms due to overdiagnosis, help in understanding the uncertainties about overdiagnosis, or both.

As a first step, we identified four questions about overdiagnosis in prostate cancer screening that modeling studies have addressed:

Questions

1. Under hypothetical screening programs and settings, how severe might the problem of overdiagnosis be?
2. How much overdiagnosis has occurred in screened populations?
3. How can screening strategies be changed to mitigate the harms of overdiagnosis?
4. How can information about possible overdiagnosis be personalized?

Next, we read the identified publications from modeling studies that reported information about overdiagnosis and categorized them according to the question(s) addressed. We found that the literature on using models to evaluate overdiagnosis for prostate cancer was easier to understand by reading the publications in chronological order by publication year, so our discussion of the studies is largely chronological. Last, we abstracted information about overdiagnosis from each publication. Because estimates of overdiagnosis are not generalizable across populations(i.e., the frequency of overdiagnosis depends on context),[16] we chose to organize the information about overdiagnosis by country or region.

Cancer Intervention and Surveillance Modeling Network (CISNET) Prostate Cancer Models

CISNET is a consortium of National Cancer Institute–sponsored investigators who conduct cancer modeling studies. The work done by these modelers has been used extensively to assess the actual and hypothetical effects of prostate cancer screening, including effects on overdiagnosis. 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 Fred Hutchinson Cancer Research Center Prostate-Specific Antigen growth and Prostate Cancer progression (PSAPC), and the University of Michigan Self-Consistency Analysis of Surveillance (SCANS) models. The findings about overdiagnosis based on these models are referred to several times in this publication. For this reason, the basic structure and main assumptions of these three CISNET prostate cancer models are shown in **Table 1**.

On the basis of the classification scheme described by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR-3), all three of the CISNET prostate cancer models would be broadly described as state-transition models[17] (i.e., the cancer process is conceptualized as a series of transitions from no cancer through cancer and death occurring over time). 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."[18] Both the PSAPC and the MISCAN-PRO models use algorithms and random draws from parametric statistical distributions to estimate outcomes. In contrast, the SCANS model represents the cancer process in terms of a series of equations that have a closed-form solution; and the model outcomes are derived analytically or numerically. In all three models, lead time is an intermediate estimate made using the model and estimates of lead time are then used to estimate the frequency of overdiagnosis.

All three CISNET prostate cancer models are quite complex and make many (and sometimes opaque) assumptions. The MISCAN-PRO model involves many parameters. Most other models used to address questions about prostate cancer overdiagnosis are not as transparent as the CISNET models, some are poorly described, and others are overly simplistic.

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

Question 1. Under Hypothetical Screening Programs and Settings, How Severe Might the Problem of Overdiagnosis Be?

Modeling studies pre-dating the results of prostate cancer screening RCTs first addressed hypothetically, how severe might the problem of overdiagnosis be. Three modeling studies[11, 16, 19] were early warnings that the magnitude of the problem of overdiagnosis for prostate cancer might be large. Two of these studies[11, 16] explored factors that affect the frequency of overdiagnosis, including possible changes in screening programs (i.e., age to start and stop screening, screening interval) that might mitigate the problem of overdiagnosis.

McGregor and colleagues[19] used a cohort model with data on prostate cancer incidence and mortality in Quebec to assess overdiagnosis defined as the proportion of screen-detected prostate cancers that would not have caused death due to cancer even if untreated (a less commonly used definition). McGregor and colleagues[19] concluded that about 84% of men with screen-detected cancer would not benefit from having the cancer detected because their cancer would not be fatal before death from causes other than prostate cancer.

Draisma and colleagues[11, 20] used the MISCAN-PRO model with data from the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC) to estimate how much overdiagnosis there might be in the Netherlands for five hypothetical screening strategies based on a PSA test threshold for biopsy referral of greater than 3 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 48% for regular screening at 55 to 67 years every 4 years (**Table 2**). Draisma and de Koning[11] thus concluded:

> "PSA advances the diagnosis of prostate cancer in time and is associated with over-detection. Studies disagree on the extent of both effects, but our results indicate that their effect could be considerable…." (p. 110)

Davidov and Zelen[16] applied the basic analytic modeling approach described by Zelen and Feinleib[7] to explore what factors might influence the frequency of prostate cancer overdiagnosis. They identified age at screening and sojourn time (i.e., the time between initiation of the cancer and when it would be clinically detected) as important influences on rates of overdiagnosis. The frequency of overdiagnosis (as a proportion of screen-detected cancer) was estimated for a range of hypothetical values of sojourn time (from 5 to 20 years) and for various hypothetical screening schedules and test sensitivities. Davidov and Zelen[16] concluded that:

> "Our calculations suggest that the probability of overdiagnosis is high and in the range of 20 to 40% for most realistic mean values [of sojourn time] and ages of screening." (p. 609)

Thus early in the prostate cancer screening discussion, Davidov and Zelen[16] note that:

"Regardless of the value of the lead time, the relatively high probability for the overdiagnosis of prostate cancer raises important issues of whether to treat a disease that has a significant probability of overdiagnosis." (p. 612)

Question 2. How Much Overdiagnosis Has Occurred in Screened Populations?

Modeling studies done before the results of RCTs were available addressed the question of how much overdiagnosis had occurred in screened populations. The answer is context-specific because the screening strategy (age to begin and end screening, and the PSA value used to guide biopsy referral), the adoption of screening, diagnostic practices (e.g., the chance that a biopsy referral will lead to biopsy, the number of biopsy cores), the incidence and natural history of prostate cancer, and life-expectancy all vary between populations. Four modeling studies[13, 21-23] attempted primarily to estimate how much of the observed increase in the incidence of prostate cancer after introduction of PSA screening in the United States from the middle to late 1980's (when PSA screening began to become widespread) through the year 2000 might be overdiagnosed. A 2002 publication[23] reported on an effort using modeling, assumptions about lead time, and data on life-expectancy to estimate the frequency of overdiagnosis as a proportion of screen-detected cancers in white and African-American [AA] men age 60 to 84 years in 1988 for the period 1988 to 1998 in United States. This study is not discussed in detail because it presents estimates of overdiagnosis only for men ages 60 years and older when screened and is described by the authors as exploratory. Information in a 2008 publication by Telesca and colleagues[21] that formalizes this modeling approach appears to supersede the findings reported in the 2002 publication.[23] The main findings about overdiagnosis based on the remaining three modeling studies are summarized in **Table 2**.

Tsodikov and colleagues[13] used an analytic model conceptually related to the model used by Zelen and Feinleib[7] and Davidov and Zelen.[16] In this model, a precursor of the SCANS model, the underlying concept of the natural history of prostate cancer consists of tumor onset, time from tumor onset to diagnosis (i.e., sojourn times), and sensitivity of the PSA test to detect latent (pre-clinical) disease. The model estimated natural history parameters based on population prostate cancer incidence rates in the period before and after widespread PSA screening, yielding estimates of the distribution of age at prostate cancer onset and sojourn time. PSA screening is superimposed to estimate lead time. Coupled with data about life expectancy derived from life tables, estimates of overdiagnosis are then calculated. Data about the uptake of PSA screening in the United States were from the National Health Interview Survey (NHIS) and linked Surveillance Epidemiology and End Results (SEER)–Medicare data. Tsodikov and colleagues[13] concluded that overdiagnosis in U.S. men screened between 1988 and 2000 depended on birth cohort and year of detection and that the frequency of overdiagnosis (as a proportion of screen-detected cancer) was higher in the early years after introduction of PSA screening but "settled in at 30% for the present [early 2000] era."

Telesca and colleagues[21] used a model that they state "formalizes the excess incidence approach" with a primary goal of estimating the lead time and its distribution for prostate cancer screening in the United States for the period after 1988 and through 2000. Modeling was used to estimate

lead time using information on population screening and disease incidence. A simulation model was then used in conjunction with the estimates of mean lead time distributions (by age) and life-expectancy to estimate frequencies of overdiagnosis (as a proportion of all screen-detected cancer) by age and race for a hypothetical cohort of 1 million men.[24] Overdiagnosis (as a proportion of screen-detected cancer) was estimated to be 23% in whites and 34% in AA men; estimates increased with age, ranging from 14% at age 50 to 54 years to 57% at age 85 years or older in AA men and 3% at age 50 to 54 years to 50% at age 85 years or older in white men.

Draisma and colleagues[22] 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 was overdiagnosed was estimated to be 8.6% (SCANS), 11.9% (PSAPC), and 18.6% (MISCAN-PRO). Draisma and colleagues[22] highlighted the importance of calibrating models using data from the population being studied and called attention to the importance of specifying definitions of lead time and overdiagnosis in publications about these topics. Their comparison of estimates of overdiagnosis from the three CISNET models helps explain the variability in previously reported estimates of overdiagnosis. On the basis of the results reported by Draisma in this comparative modeling exercise, it is reasonable to conclude that both the proportions of screen-detected prostate cancers and all prostate cancers that were overdiagnosed in the United States during the late 1980's through 2000 were sizable.

Question 3. How Might Screening Schedules and Strategies Be Changed to Mitigate Harms of Overdiagnosis?

A question commonly addressed in modeling studies to address overdiagnosis is how screening schedules and strategies can be changed to mitigate the harms of overdiagnosis. Because estimates of overdiagnosis depend on life expectancy, such studies need to be done with data about life expectancy from the population in which the screening schedule would be implemented. As shown by Draisma and de Koning,[11] the model should be calibrated to data from the population to which the information will be applied. Modeling studies that address this question have been done for the United States[25, 26] and Canada[27] using the PSAPC model, for the United States using the MISCAN-PRO model,[28] for the Netherlands/Western Europe using the MISCAN-PRO model[20, 29-31] for Finland,[32] and for the United Kingdom using different multi-state Markov models.[33] **Table 2** summarizes the major findings of these studies; however, we discuss in detail only publications reporting data for the United States or the Netherlands/Western Europe based on CISNET models.

United States

Three main modeling studies pertinent to the United States reported on overdiagnosis for various screening strategies.[25, 26, 28] Gulati and colleagues[25] used the PSAPC model to evaluate four biennial screening strategies—age 50 to 74 years or 50 to 84 years—with PSA thresholds for

biopsy referral of greater than 4 µg/L or 2.5 µg/L. Among the four modeled screening strategies, Gulati and colleagues[25] concluded that screening with a PSA cut-point of 4.0 µg/L and screening ages 50 to 74 years "performs best in terms of overdiagnoses per prostate cancer detected." Because the main aim of this study was to develop and calibrate the PSAPC model, the authors clearly identified their results as non-comprehensive.

A subsequent publication based on the PSAPC model[26] reported the absolute risk of overdiagnosis (as a proportion of men screened, referred to as the "probability of overdiagnosis" in the publication) for a more comprehensive set of 35 screening strategies with various stop and start ages, screening intervals, and PSA criteria for biopsy referral. Also presented in this publication were estimates of other harms of screening (false-positive results) and benefits of screening (probability of cancer death, mean time of life saved), which we do not discuss here. **Table 3** provides estimates of the lifetime risk of overdiagnosis (overdiagnosed cancer as a proportion of men screened) for the eight strategies that evaluated a single value of PSA as a threshold for biopsy referral (as opposed to an age-dependent PSA threshold or a threshold based on PSA velocity) and for one complex strategy (strategy 1, the National Comprehensive Cancer Network [NCCN] strategy) that involved tailored intervals and thresholds based on age as well as variable PSA criteria for biopsy. Selected data are presented, as in the original publication, in order of increasing probability of prostate cancer death; the data in **Table 3** covers the range of estimates of overdiagnosis presented in the publication.[26] The estimates of the lifetime risk of overdiagnosis ranged from 6.0% for the complex NCCN strategy to 1.3% for biennial screening of men 50 to 69 years old with a PSA threshold of 4.0 µg/L. The authors concluded that:

> "…more aggressive screening strategies, particularly those that lower the PSA threshold for biopsy, do reduce mortality relative to the reference strategy. However, the harms of unnecessary biopsies, diagnoses, and treatments may be unacceptable." (p. 151-152) [26]

The authors emphasized that less aggressive strategies should involve weighing the harms, including overdiagnosis, against possible benefits. This modeling study did not produce estimates of quality-adjusted life years, as the authors acknowledge the limited data and reliability of data on utilities associated with prostate cancer screening and postdiagnosis health states.

Using the MISCAN-PRO model, de Carvalho and colleagues[28] published estimates of overdiagnosis and prostate cancer harms for an expanded set of screening strategies for the United States. As in Gulati and colleagues' work,[26] changes in hypothetical screening strategies affected the estimates of overdiagnosis. The lowest estimate of the lifetime risk of overdiagnosis (as a proportion of men screened) was 1.99% for a strategy in which men age 50 to 70 years old were screened every 4 years with a PSA threshold for biopsy referral greater than 3 µg/L. The highest estimate of the lifetime risk for overdiagnosis was 6.26% for a strategy in which men 50-80 years were screened annually with a PSA threshold for biopsy referral greater than 3 µg/L.

Netherlands/Western Europe

A series of four publications that evaluated various screening strategies with the use of the MISCAN-PRO model reported estimates of overdiagnosis pertinent to Western Europe.[20, 29-31].

In 2003, Draisma and colleagues[20] reported on five screening schedules, different from those addressed in 2009[30] and 2012.[31] Estimates of overdiagnosis based on the MISCAN-PRO model presented in a 2015 publication by Heijnsdijk and colleagues[29] appeared to be identical to estimates in the 2012 publication.[31] Unlike the modeling studies pertinent to the United States that were previously discussed, the metric used for overdiagnosis in these publications was overdiagnosed cancer as a proportion of screen-detected cancer. Also, in these studies pertinent to Western Europe a much smaller number of potential prostate cancer screening strategies are explored, as compared to the modeling studies pertinent to the United States. The findings are summarized in **Table 4**.

As expected, estimates of overdiagnosis depend on age and screening interval, such that screening at older ages was associated with a higher estimated overdiagnosis and a longer interval between screening was associated with a lower estimated frequency of overdiagnosis. Across these four MISCAN-PRO publications, the estimates of overdiagnosis (as a proportion of screen-detected cancer) for various schedules ranged from a low of 27%[20] to 30%[29, 31] for a single screen at age 55 years to a high of 57%[30] for screening from age 55 to 75 years every 4 years. The estimates of overdiagnosis were more than 50% for all schedules that involved screening men at age 75 years once, or included screening men age 75 years or older. It is noteworthy that the 2012 estimates of overdiagnosis based on the MISCAN-PRO model did not evaluate any schedules that included screening men aged 75 years.[31]

Question 4. How Can Information About (Possible) Overdiagnosis Be Personalized?

Three modeling studies[34-36] primarily addressed how information about possible overdiagnosis be personalized. Those studies evaluated clinical factors (e.g., age, clinical stage, PSA, Gleason score, co-morbidities) that might affect prognosis and overdiagnosis, such that their findings could conceivably be used to counsel men with screen-detected prostate cancer about the probability of overdiagnosis such that men at high overdiagnosis risk might reasonably choose to pursue active surveillance as opposed to immediate radical therapy (i.e., prostatectomy or radiation). However, the use of the data from the modeling studies to select men who would enter active surveillance in clinical populations has not yet been evaluated.

Wever and collagues[34] used the MISCAN-PRO model to estimate the chances of overdiagnosis in men 55 to 74 years old with prostate cancer detected by PSA screening according to factors that might affect prognosis (i.e., age, clinical stage, and Gleason score). Estimates of benefits based on prognostic factors were also presented but we do not discuss them here. The model used 2000 to 2007 Netherlands life tables to estimate life expectancy, so the findings would likely apply to countries with a preponderantly white population and similar life expectancy. As discussed previously, the context of the population affects the estimates of overdiagnosis, and European pertinent model estimates of overdiagnosis appear to be higher than United States pertinent model estimates. Estimates of overdiagnosis (as a proportion of screen-detected cancer) showed a wide range of values: from 2.7% for a man aged 55 to 59 years with clinical stage T3 and high-grade cancer (Gleason score ≥7) to 60.1% for a man aged 70 to 74 years with a clinical stage T1 and lower-grade cancer (Gleason score <7). Estimated overdiagnosis was related to age

(i.e., the frequency of overdiagnosed cancer as a proportion of screen-detected cancers increases with age within clinical stages and categories of Gleason score), clinical stage (i.e., the estimated frequency of overdiagnosed cancer as a proportion of screen-detected cancers is higher at clinical stage T1 than at stage T2 which is higher than at stage T3 within age bands and categories of Gleason score), and Gleason score (i.e., the estimated frequency of overdiagnosis as a proportion of screen-detected cancer is lower in higher categories of Gleason score within clinical stage at the same age).

In another modeling study with similar aims, Gulati and collagues[35] used the PSAPC model to derive individualized estimates of the likelihood that a screen-detected cancer would be overdiagnosed based on age, PSA level, and Gleason score. They presented a nomogram based on the model that could be used to estimate the lifetime risk of overdiagnosis as a percentage for a man with screen-detected cancer considering age, PSA level, and Gleason score. Illustrative risks of overdiagnosis in this study showed a wide range of values. For example, the estimated risk of overdiagnosis is 5.0% for prostate cancer detected at age 50 to 54 years with a Gleason score of less than 7 and a PSA level of 9.0 to 9.9 µg/L, but 83.4% for prostate cancer detected at age 80 to 84 years with a Gleason score less than 7 and a PSA level of 4.0 to 4.9 µg/L.

Lansdorp-Vogelaar and collagues[36] reported on an effort in which breast, prostate, and colorectal cancer CISNET modelers collaborated to derive estimates of the chances of overdiagnosis with the aim of using the information to personalize decisions about the age to end screening based on co-morbidity for individuals over age 65 years. Data on benefits (i.e., cancer death prevented and quality-adjusted life-years [QALYs] gained) and other harms (i.e., false-positive test results) were also reported but are not discussed here. The lifetime risk of overdiagnosis (i.e., as a proportion of screened individuals) in categories of co-morbidity (no co-morbidity, mild co-morbidity, moderate co-morbidity, severe co-morbidity) was estimated for the U.S. population aged 66 to 90 years in 2010 who continued to be screened regularly compared with stopping screening at various modeled ages. Estimates were made using PSAPC and MISCAN-PRO for prostate cancer, MISCAN-Fadia and Georgetown-Einstein for breast cancer, and MISCAN-Colon, Colorectal Cancer Simulated Population model for Incidence and National history (CRC SPIN), and Simulating Colorectal Cancer (SimCRC) for colorectal cancer. The investigators assumed that all persons aged 65 years or older had had regular screening starting at age 50 years with biennial PSA testing (threshold not specified), biennial mammography, and annual fecal immunochemical testing (FIT). All models were calibrated to data from the SEER program.

Summarized data on prostate cancer overdiagnosis by Lansdorp-Vogelaar and collagues[36] are provided in **Table 5**. For both prostate cancer models, within categories of age at screening, the estimated absolute risk of overdiagnosis (overdiagnosed cancer as a proportion of men screened) increased with increasing co-morbidity, which was not surprising because co-morbidity decreases life-expectancy. The estimates of the absolute risk of overdiagnosis were generally 0.5 to 1 percentage points higher for the MISCAN-PRO model than those based on the PSAPC model.

A comparison of the estimates of overdiagnoses across the three types of cancer illustrates the clinical significance of overdiagnosis in prostate cancer screening. In individuals with "average" comorbidity (and thus average life expectancy) screened for prostate cancer at age 74 years, the

absolute risk of overdiagnosis (as a proportion of individuals screened) was estimated to be 1.97% based on the MISCAN-PRO model and 1.45% based on the PSAPC model. In contrast, in individuals with "average" comorbidity screened for breast cancer at age 74 years, the absolute risk of overdiagnosis was estimated to be 0.08% based on the MISCAN-Fadia model and 0.05% based on Georgetown-Einstein model. In individuals with "average" comorbidity screened for colorectal cancer at age 74 years, the absolute risk of overdiagnosis was estimated to be 0.03% based on the MISCAN-Colon model, 0.00% based on the CRC-SPIN model, and 0.01% based on the SimCRC model. Comparing prostate, breast and colorectal cancer screening, Lansdorp-Vogelaar and collagues[36] state that:

> "the balance of benefits and harms…were mostly similar except for the rates of overdiagnosis, which were orders of magnitude (15 to >100 times, depending on the model) greater for prostate cancer versus breast or colorectal cancer screening." (p. 107)

Chapter 4. Discussion

Critical Assessment of Model-Based Information on Overdiagnosis

Model-based estimates of overdiagnosis have used an array of numerator and denominator definitions. The exact measures being reported can be difficult to glean from a publication.[10] Related to the issue of metrics, in the mid-2000's, modeling studies evaluating the effect of screening strategies that were meant to reflect the United States experience began to report overdiagnosis as a proportion of the number of individuals screened (i.e., the absolute risk) instead of overdiagnosis as a proportion of screened-detected cancers. Which (if any) metric—overdiagnosis as a proportion of screen-detected cancers, as a proportion of all cancers, as an absolute risk, or as something else—is most meaningful to patients, physicians, and policy-makers appears not to have been assessed using empiric data.

An additional problem with reading the modeling literature arises because there are often several publications that report data about overdiagnosis based on the same model and/or meant to apply to the same population. It is difficult to know whether a new set of results supersedes or supplements an older set of results. It is also difficult to identify changes in model assumptions between uses of the model to assess overdiagnosis that might invalidate prior results or explain discrepancies between the estimates.

The existence of several publications from the same study is a well-known issue for systematic reviews, but picking one publication to represent "the" model estimate of overdiagnosis is not a good solution for the modeling literature because models evolve and the same models are used to address new questions. Publications about overdiagnosis based on modeling should provide specific guidance to reconcile the results of studies based on the same model and/or conducted by the same modeling group and/or addressing the same question.

Summary and Conclusions

The model-based literature for prostate cancer screening addressing overdiagnosis is abstruse. Given the potential implications for policy decisions regarding prostate cancer screening, it is necessary to understand the findings and limitations of the models, as both contribute to efforts to assess the magnitude and certainty of estimates of overdiagnosis. The model-based literature to address overdiagnosis, as well as the model's assumptions, have necessarily evolved over time; however these changes add complexity to understanding differences or discrepancies between model findings over time.

When discussing overdiagnosis, it is crucial to be specific about how overdiagnosis is being defined and to whom it applies. The modeling studies we identified use an array of definitions and metrics to capture overdiagnosis, this is sometimes not transparent, and it creates difficulty in comparing estimates across modeling studies, as well as between modeling and empiric

studies. Estimates of overdiagnosis are dependent on many factors which vary across different populations (e.g., natural history of cancer, life expectancy of the population, differences in clinical practice [e.g., PSA threshold for biopsy referral, adherence with biopsy recommendation, number of biopsy cores], and prior history of screening). As such, estimates of overdiagnosis, no matter what metric is being used, vary across populations and over time.

The major harm of screening, overdiagnosis is the key factor in determining the net benefit of prostate cancer screening. Modeling studies suggest that the frequency of overdiagnosis, by all metrics, is much greater for prostate cancer screening than for breast or colorectal cancer screening and that the proportion of screen-detected prostate cancers among men screened from the late 1980s and 2000 that were overdiagnosed was sizable. Because empirical data to estimate overdiagnosis is quite limited for prostate cancer screening, modeling data can be helpful in bounding estimates of overdiagnosis and exploring how altering or targeting screening strategies could mitigate the frequency of overdiagnosis. Although the modeling studies suggested that raising the PSA threshold, lengthening the interval of screening, and lowering the age to stop screening would all reduce the frequency of overdiagnosis, no screening strategy can eliminate the problem of overdiagnosis. Clinical factors that affect overdiagnosis such as age, PSA level, clinical stage of the cancer, Gleason score, and patient's comorbidities, could all potentially be used to counsel men with screen-detected prostate cancer about the chance that the cancer would have not otherwise been clinically detected in their lifetime, permitting more informed choices about treatment. A modeling study by Roth and colleagues[37] discussed in a companion paper[38] suggested that screening strategies that involve less aggressive management of screen-detected cancer—greater use of active surveillance in low-risk cancers— might preserve the benefit of screening while reducing harms. The use of the data from modeling studies to select men who could enter active surveillance in clinical populations has not yet been evaluated.

Overtreatment is the major potential harm of overdiagnosis. However, even with reduction of potential treatment harms through active surveillance, overdiagnosis can result in psychosocial morbidity, a burden of testing, and the diagnosis of cancer may have financial implications even beyond increased costs of medical care. Given the range of potential sequela from overdiagnosis, the true disutility of overdiagnosis is not known and has not been measured empirically. Therefore, the balance of benefits to harms is, in part, dependent on the (differences) in preferences and values related to avoiding overdiagnosis.

References

1. Love RR. The efficacy of screening for carcinoma of the prostate by digital examination. American journal of preventive medicine. 1985;1(6):36-46. PMID: 3916716.
2. Chou R, Dana T, Bougatsos C, et al. Treatments for Localized Prostate Cancer: Systematic Review to Update the 2002 U.S. Preventive Services Task Force Recommendation. Rockville (MD): Agency for Healthcare Research and Quality; 2011. PMID: 22171386.
3. Donovan JL, Hamdy FC, Lane JA, et al. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med. 2016;375(15):1425-37. PMID: 27626365. https://dx.doi.org/10.1056/NEJMoa1606221
4. Brawley OW. Trends in prostate cancer in the United States. J Natl Cancer Inst Monogr. 2012;2012(45):152-6. PMID: 23271766. https://dx.doi.org/10.1093/jncimonographs/lgs035
5. Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986-2005. J Natl Cancer Inst. 2009;101(19):1325-9. PMID: 19720969. https://dx.doi.org/10.1093/jnci/djp278
6. Moller MH, Kristiansen IS, Beisland C, et al. Trends in stage-specific incidence of prostate cancer in Norway, 1980-2010: a population-based study. BJU Int. 2016;118(4):547-55. PMID: 26497872. https://dx.doi.org/10.1111/bju.13364
7. Zelen M, Feinleib M. On the theory of screening for chronic diseases. Biometrika. 1969;56(3):601-14. https://doi.org/10.1093/biomet/56.3.601
8. Walter SD, Day NE. Estimation of the duration of a pre-clinical disease state using screening data. Am J Epidemiol. 1983;118(6):865-86. PMID: 6650488.
9. Etzioni R, Gulati R, Mallinger L, et al. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Ann Intern Med. 2013;158(11):831-8. PMID: 23732716. https://dx.doi.org/10.7326/0003-4819-158-11-201306040-00008
10. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ (Clinical research ed). 2015;350:g7773. PMID: 25569206. https://dx.doi.org/10.1136/bmj.g7773
11. 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
12. World Health Organization. Life expectancy, Data by country. http://apps.who.int/gho/data/node.main.688?lang=en. Accessed May 3, 2017. PMID.
13. 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
14. Gulati R, Feuer EJ, Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies. American journal of epidemiology. 2016;184(2):140-7. PMID: 27358266. https://dx.doi.org/10.1093/aje/kwv342
15. Etzioni R, Gulati R. Oversimplifying overdiagnosis. Journal of general internal medicine. 2014;29(9):1218-20. PMID: 24830739. https://dx.doi.org/10.1007/s11606-014-2867-0
16. 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

17. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2012;15(6):812-20. PMID: 22999130. https://dx.doi.org/10.1016/j.jval.2012.06.014
18. Cancer Intervention and Surveillance Modeling Network. Prostate Cancer Modeling. https://cisnet.cancer.gov/prostate/. Accessed May 3, 2017.
19. 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.
20. 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.
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, 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
23. 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
24. Mariotto AB, Etzioni R, Krapcho M, et al. Reconstructing PSA testing patterns between black and white men in the US from Medicare claims and the National Health Interview Survey. Cancer. 2007;109(9):1877-86. PMID: 17372918. https://dx.doi.org/10.1002/cncr.22607
25. 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
26. 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
27. 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
28. 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
29. 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
30. 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

31. 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
32. 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
33. 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
34. 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
35. 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
36. 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
37. Roth JA, Gulati R, Gore JL, et al. Improving the value of PSA prostate cancer screening with "smarter" strategies and increased use of active surveillance (2015 ASCO Annual Meeting May 29-June 2, 2015). J Clin Oncol. 2015;33(Abstract 6616):S15, May 20 Supplement.
38. Lin JS, Petitti DB, Burda BU. Contextual Review: Overview of 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.

Table 1. 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 1. Overview of CISNET Prostate Cancer Screening Models

Model Institution Type	Population Main Data Sources	Inputs and Assumptions Natural History	Screening	Treatment	Outcomes
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 depend 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 andobservational 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 1. 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 2. Summary of Prostate Cancer Screening Modeling Studies of Overdiagnosis by Questions Addressed

Author, Year	Purpose of Modeling Study	Model(s) Population	Definition and Metric to Measure Overdiagnosis	Study Findings
Question 1. Under hypothetical screening programs and settings, how severe might the problem of overdiagnosis be?				
McGregor et al, 1998[19]	Estimate amount of overdiagnosis that might result from annual PSA screening of men age 50-70 years with PSA threshold >4 at "steady state"	Macrosimulation/ outcomes table Quebec population, mortality from Quebec Vital Statistics 1988-1992	Proportion of screen-detected cancers that would not have resulted in death due to cancer even if treated ("nonstandard definition")	An estimated 84% (78%-87%) of men would be considered overdiagnosed Assuming that radical prostatectomy is 100% effective, only 16 (at most 22) out of 100 men undergoing surgery could likely benefit from the surgery.
Draisma et al, 2003[11, 20]	Estimate mean lead time and rate of overdiagnosis associated with PSA screening for 5 schedules	MISCAN-PRO* European population, mortality from life table statistics Netherlands 1991-1995	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	The frequency of overdiagnosis depends on schedule; for example, the estimated frequency of overdiagnosis (1) is[§]: 27% (24%-37%) single screen at age 55 years 38% (34%-47%) single screen at age 60 years 47% (43%-55%) single screen at age 65 years 53% (50%-60%) single screen at age 70 years 56% (53%-61%) single screen at age 75 years 50% (46%-57%) annual screen age 55-67 years 56% (54%-61%) annual screen age 55-75 years 48% (44%-55%) screen every 4 years age 55-67 years 54% (51%-59%) screen every 4 years age 55-75 years
Davidov and Zelen, 2004[16]	Analyze an idealized early detection program and derive the mathematical expression for the probability of overdiagnosis	Three-stage process based on chronic disease model US population, mortality from SSA period life-table 1997	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	Frequency of overdiagnosis depends on sojourn time (increases with longer sojourn times), age (increases with age for all sojourn times), and screening schedule (increases with repeated screening). Overall the probability of overdiagnosis (1) is substantial and ranges from 20%-40% for realistic estimates of sojourn time and ages of screening.
Question 2. How much overdiagnosis has occurred in screened populations?				
Tsodikov et al, 2006[13]	Estimate lead time and overdiagnosis in US men after dissemination of PSA screening through 2000	SCANS* US population, mortality from Human Mortality Database, years NR	Proportion (1) of screen-detected cancers, or (2) of all detected cancers that would not have otherwise been clinically detected during patient's lifetime	The estimated frequency of overdiagnosis from men screened from 1988 to 2000 in the US depends on birth cohort and age at detection. The estimated frequency of overdiagnosis changed over time and "settled in" at 30% when defined as (1) a proportion of screen-detected cancers and 25% when defined (2) as a proportion of all detected cancers.

Table 2. Summary of Prostate Cancer Screening Modeling Studies of Overdiagnosis by Questions Addressed

Author, Year	Purpose of Modeling Study	Model(s) Population	Definition and Metric to Measure Overdiagnosis	Study Findings
Telesca et al, 2008[21]	Estimate lead time and overdiagnosis in US men after dissemination of PSA screening through 2000	Analytic model that does not conceptualize natural history of cancer, rather the effects of screening on incidence US population mortality based on US mortality from 1992 CDC National Center for Health Statistics Vitals Statistics	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	Overdiagnosis (1) was estimated to be 23% in whites and 34% in AA men. Estimates increased with age, ranging from 14% at age 50-54 years to 57% at age 85 years or older in AA men, and 3% at age 50-54 years to 50% at age 85 years or older in white men.
Draisma et al, 2009[22]	Estimate overdiagnosis in US men from 1985-2000 comparing estimates between CISNET models	PSAPC, MISCAN-PRO, and SCANS* US population for all models, mortality from standard US life tables, years NR	Proportion (1) of screen-detected cancers, or (2) of all detected cancers that would not have otherwise been clinically detected during patient's lifetime	Estimates of the frequency of overdiagnosis in the US 1985-2000 differ between models. Defined as (1) a proportion of screen-detected cancer, the estimated frequency of overdiagnosis was: 28.0% PSAPC 22.9% SCANS 42.0% MISCAN-PRO Defined as (2) a proportion of all cancer, the estimated frequency of overdiagnosis was: 11.9% PSAPC 8.6% SCANS 18.6% MISCAN-PRO
Question 3. How might screening schedules and strategies be changed to mitigate harms of overdiagnosis?				
Gulati et al, 2010[25]	Project early detections, overdiagnoses, and mean lead times for 4 candidate screening strategies: age 50-74 or 50-84 with PSA cut-off >4.0 µg/L or >2.5 µg/L (with biennial screening interval)	PSAPC* US population, mortality from standard US life tables, years NR	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	Screening with a PSA cut-point of 4.0 vs. 2.5 µg/L, performed better in terms of overdiagnosis. The estimated overdiagnosis (1) is: for age 50-84 y, ~40% (cut-point 2.5 µg/L) vs. ~35% (cut-point 4.0 µg/L) for age 50-74 y, ~27% (cut-point 2.5 µg/L) vs. ~23% (cut-point 4.0 µg/L)
Gulati et al, 2013[26]	Comparative effectiveness of alternative PSA screening strategies. Explores 35 strategies that vary by start and stop age, screening interval, and threshold for biopsy referral	PSAPC* US population, mortality from standard US life tables, years NR	Proportion (3) of men screened that would not have otherwise been clinically detected during patient's lifetime	Estimates of overdiagnosis (3) ranged from 1.3% to 6.0% depending on screening strategy‖

Table 2. Summary of Prostate Cancer Screening Modeling Studies of Overdiagnosis by Questions Addressed

Author, Year	Purpose of Modeling Study	Model(s) Population	Definition and Metric to Measure Overdiagnosis	Study Findings
de Carvalho et al, 2015[28]	Comparative effectiveness of alternative PSA screening strategies. Adds more screening strategies to those assessed in Gulati et al, 2013.	MISCAN-PRO* US population, mortality from standard US life tables, years NR	Proportion (3) of men screened that would not have otherwise been clinically detected during patient's lifetime	Estimates of overdiagnosis (3) ranged from 1.99% to 6.26% depending on screening strategy
Heijnsdijk et al, 2009[30]	Simulate effects of 3 different prostate cancer screening programs using a PSA cut-point of 3.0 ng/dL, in hypothetical population of 100,000 men	MISCAN-PRO* European population, mortality from European Standard population 2003	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	Estimates of of overdiagnosis (1) are[§]: 39% screen age 55-70, every year 40% screen age 55-70, every 2 years 51% screen age 55-75, every 4 years
Heijnsdijk et al, 2012[31]	Predict prostate cancers, treatments, deaths and QALYs for 6 screening schedules using a PSA cut-point of 3.0 ng/dL	MISCAN-PRO* European population, mortality from European Standard population, years NR	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	Estimates of overdiagnosis (1) are[§]: 30% screen age 55, once 35% screen age 60, once 45% screen age 65, once 43% screen age 55-69, every year 48% screen age 55-74, every year 41% screen age 55-69, every 4 years
Pataky et al, 2014[27]	Comparative effectiveness of 14 different screening strategies using a PSA cut-point of 3.0 or 4.0 ng/dL, includes QALYs	PSAPC* British Columbia, source for mortality NR	Proportion (2) of all detected cancers that would not have otherwise been clinically detected during patient's lifetime	Estimates overdiagnosis (2) range from 0.06% to 23.1%. Selected estimates are: 0.06% screen age 50, one-time screen 8.4% screen age 55-69, every 4 years 20.7% screen age 60-74, every 2 years 16.1% screen age 50-74 years, every 2 years[‡] 21.9% screen age 40-74 years, every 2 years
Question 4. How can information about (possible) overdiagnosis be personalized?				
Wever et al, 2013[34]	Estimate benefits and harms of prostate cancer screening in men 55-74 years by prognostic factors (age, stage, and Gleason score)	MISCAN-PRO* European population, mortality from life-table statistics Netherlands 2000-2007	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	The estimates of overdiagnosis (1) have a wide range of values. The lowest and highest estimates of overdiagnosis are: 2.7% clinical stage T3, Gleason score >7, age 55-59 years 60.1% clinical stage T1, Gleason score <7, age 70-74 years Overdiagnosis (1) increases with age within clinical stages and categories of Gleason score. At the same age and category of Gleason score, overdiagnosis (1) is higher at clinical stage T1 than at stage T2, which is higher than at stage T3. At the same age and clinical stage, overdiagnosis (1) is lower in higher categories of Gleason score

Table 2. Summary of Prostate Cancer Screening Modeling Studies of Overdiagnosis by Questions Addressed

Author, Year	Purpose of Modeling Study	Model(s) Population	Definition and Metric to Measure Overdiagnosis	Study Findings
Gulati et al, 2014[35]	Derive individualized estimates of overdiagnosis by prognostic factors (age, PSA, and Gleason score)	PSAPC* US population, mortality from standard US life tables, years NR	Proportion (1) of screen-detected cancers that would not have otherwise been clinically detected during patient's lifetime	Estimates of overdiagnosis (1) have a wide range of values. The highest and lowest estimates of overdiagnosis (1) are: 4.1% age 50-54 years, Gleason score ≥7, PSA level 9.0-9.9 µg/L 83.4% age 80-84 years, Gleason score ≤6, PSA level 4.0-4.9 µg/L Overdiagnosis (1) increases with age within category of Gleason score and PSA. At the same age and category of Gleason score, overdiagnosis (1) decreases with higher PSA level. At the same age and category of PSA, overdiagnosis (1) is lower at Gleason score ≥7 than at Gleason score ≤6
Lansdorp-Vogelaar et al, 2014[36]	Derive estimates of benefits and harms of screening that would permit personalization of age to end screening based on comorbidity in individuals age >65 years	PSAPC and MISCAN-PRO*† US population, mortality from Medicare data from SEER areas, 1992-2005	Proportion (3) of persons screened that would not have otherwise been clinically detected during patient's lifetime	For prostate cancer, within categories of age at screening, overdiagnosis (3) generally increases with increasing comorbidity.¶ Depending on the model used, estimates of overdiagnosis (3) were "orders of magnitude (15 to 100 times)" greater for prostate cancer screening vs. breast or colorectal cancer screening in the population age 65 years or older.

*See Table 1
†Other CISNET models were used to assess breast cancer and colorectal cancer; all were for the U.S. population
‡PSA threshold of 3.0 µg/L up to age 69 years, then 4.0 µg/L at age 70 years or older
§See Table 4
‖See Table 3
¶See Table 5

Abbreviations: AA = African-American; FHCRC = Fred Hutchinson Cancer Research Center; MISCAN-PRO = Microsimuliaton Screening Analysis Prostate Cancer model; NR = not reported; PSA = prostate-specific antigen; PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression; SCANS = Self-Consistency Analysis of Surveillance; SSA = Social Security Administration

Table 3. Lifetime Risk of Overdiagnosis (Overdiagnosed Cancer as a Proportion of Men Screened) by Screening Schedule or Strategy for the United States Using PSAPC (Ordered According to Estimated Prevented Prostate Cancer Deaths per 100 Men Regularly Screened)[26]

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, %
1 (based on NCCN)	40-74	Complex†	Complex‡	6.0	2.02
5	50-74	Annual	>2.5	4.7	2.08
8	50-74	Annual	>4.0	3.3	2.15
11	50-74	Biennial	>2.5	3.8	2.16
18	50-74	Biennial	>4.0	2.7	2.23
21	50-69	Annual	>2.5	2.9	2.24
26	50-69	Annual	>4.0	1.8	2.32
30	50-69	Biennial	>2.5	2.0	2.35
35	50-69	Biennial	>4.0	1.3	2.43

†Annual (quinquennial if age <50 and PSA <1 µg/L)
‡PSA >2.5 µg/L or PSA velocity >0.35 µg/L per year

Abbreviations: PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression; NCCN = National Comprehensive Cancer Network; PSA = prostate-specific antigen

Table 4. Lifetime Risk of Overdiagnosis (Overdiagnosed Cancer as a Proportion of Screen-Detected Cancers) by Screening Schedule/Strategy for Netherlands/Europe Using MISCAN-PRO

Age, years	Screening Interval	Estimated Overdiagnosis as a Proportion of Screen-Detected Cancers (Range)*		
		Draisma, 2003[11, 20]	Heijnsdijk, 2009[30§†]	Heijnsdijk, 2012[31] and 2015[29‖†]
55	Once	27 (24-37)	--	30 (NR)
60	Once	38 (34-47)	--	35 (NR)
70	Once	53 (50-60)	--	45 (NR)
75	Once	56 (53-61)	--	--
55-67	Annually	50 (46-57)	--	--
55-69	Annually	--	--	43 (NR)
55-70	Annually	--	49 (NR)	--
55-74	Annually	--	--	48 (NR)
55-75	Annually	56 (54-61)	--	--
55-70	Biennially	--	48 (NR)	--
55-67	Every 4 years	48 (44-55)	--	41 (NR)
55-70	Every 4 years	--	42 (NR)¶	--
55-75	Every 4 years	54 (51-59)	57 (NR)	--

*PSA greater than 3µ/L
†Assume 100% participation
§European standard population 2003 for screening 2003-2033
‖Assumes 80% participation in screening program
¶Base program

Abbreviation: MISCAN-PRO = Microsimulation Screening Analysis Prostate Cancer; NR = nor reported

Table 5. Lifetime Risk of Overdiagnosis (Overdiagnosed Cancer as a Proportion of Men Screened) and Cancer Deaths Prevented in Hypothetical Scenario of Biennial Screening for Prostate Cancer by Age and Comorbidity for the United States Using MISCAN-PRO and PSAPC[36]

Age at Screen, years	MISCAN-PRO			PSAPC		
	Estimated Overdiagnosis as a Proportion of Individuals Screened	Estimated Prevented Prostate Cancer Deaths per 100 Regularly Screened	Estimated Life-Years Gained per 100 Regularly Screened	Estimated Overdiagnosis as a Proportion of Individuals Screened	Estimated Prevented Prostate Cancer Deaths per 100 Regularly Screened	Estimated Life-Years Gained per 100 Regularly Screened
No comorbidity						
66	0.86%	0.11	0.87	0.62%	0.10	1.10
74	1.84%	0.14	0.78	1.22%	0.10	0.84
80	2.99%	0.11	0.50	1.73%	0.09	0.51
86	3.25%	0.05	0.15	2.24%	0.06	0.25
90	2.83%	0.02	0.05	2.54%	0.05	0.14
Mild comorbidity						
66	0.92%	0.10	0.79	0.69%	0.09	1.01
74	1.90%	0.13	0.70	1.27%	0.10	0.77
80	3.07%	0.10	0.44	1.80%	0.08	0.45
86	3.30%	0.04	0.13	2.28%	0.06	0.23
90	2.83%	0.02	0.05	2.54%	0.04	0.12
Moderate comorbidity						
66	1.00%	0.09	0.68	0.79%	0.08	0.86
74	2.07%	0.10	0.54	1.44%	0.08	0.60
80	3.29%	0.08	0.29	1.95%	0.06	0.31
86	3.43%	0.03	0.10	2.38%	0.05	0.17
90	2.98%	0.01	0.03	2.68%	0.03	0.09
Severe comorbidity						
66	1.23%	0.06	0.45	1.02%	0.06	0.61
74	2.39%	0.06	0.33	1.66%	0.06	0.41
80	3.62%	0.05	0.18	2.13%	0.04	0.21
86	3.66%	0.02	0.06	2.50%	0.03	0.12
90	3.07%	0.01	0.02	2.76%	0.02	0.07
Average comorbidity*						
74	1.97%	0.12	0.66	1.45%	0.08	0.61
76	2.45%	0.12	0.63	1.62%	0.07	0.51

Estimated absolute risks of overdiagnosis were converted to a denominator of 100 for ease of comparison across modeling studies from a denominator of 1000 used in the original publication

*Average co-morbidity or all co-morbidities

Abbreviations: MISCAN-PRO = Microsimulation Screening Analysis Prostate Cancer; PSAPC = Prostate-Specific Antigen Growth and Prostate Cancer Progression