

UCSF

UC San Francisco Previously Published Works

Title

Limitations of Basing Screening Policies on Screening Trials

Permalink

<https://escholarship.org/uc/item/68h08673>

Journal

Medical Care, 51(4)

ISSN

0025-7079

Authors

Etzioni, Ruth
Gulati, Roman
Cooperberg, Matt R
et al.

Publication Date

2013-04-01

DOI

10.1097/mlr.0b013e31827da979

Peer reviewed



Published in final edited form as:

Med Care. 2013 April ; 51(4): 295–300. doi:10.1097/MLR.0b013e31827da979.

Limitations of basing screening policies on screening trials: The US Preventive Services Task Force and prostate cancer screening

Ruth Etzioni, PhD,

Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-B230, Seattle WA 98109-1024, Tel: +1.206.667.6561, Fax: +1.206.667.7264, retzioni@fhcrc.org

Roman Gulati, MS,

Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-B230, Seattle WA 98109-1024, Tel: +1.206.667.7795, Fax: +1.206.667.7264, rgulati@fhcrc.org

Matt R Cooperberg, MD,

Department of Urology, UCSF Helen Diller Family Comprehensive Cancer Center, San Francisco CA 94143-1695, Tel: +1.415.885.3660, Fax: +1.415.885.7443, mcooperberg@urology.ucsf.edu

David M Penson, MD,

Vanderbilt University Medical Center, 2525 West End Avenue, Suite 600, Nashville TN 37203-1738, Tel: +1.615.343.1529, Fax: +1.615.321.6350, david.penson@vanderbilt.edu

Noel S Weiss, PhD, and

Department of Epidemiology, University of Washington, 1959 NE Pacific Street, Health Sciences F-262D, Seattle WA 98195, Tel: +1.206.685.1788, Fax: +1.206.543.8525, nweiss@u.washington.edu

Ian M Thompson, MD

Department of Urology, University of Texas, 7703 Floyd Curl Drive, San Antonio TX 78229-3900; Tel: +1.210.567.5643; Fax: +1.210.567.6868; thompsoni@uthscsa.edu

Abstract

Background—The US Preventive Services Task Force recently recommended against prostate-specific antigen (PSA) screening for prostate cancer based primarily on evidence from the European Randomized Study of Screening for Prostate Cancer (ERSPC) and the US Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer screening trial.

Objective—To examine limitations of basing screening policy on evidence from screening trials.

Methods—We review published modeling studies that examine population and trial data. The studies (1) project the roles of screening and changes in primary treatment in the US mortality decline, (2) extrapolate the ERSPC mortality reduction to the long-term US setting, (3) estimate

Name and address for correspondence and reprint requests: Ruth Etzioni, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-B230, Seattle WA 98109-1024; Tel: +1.206.667.6561; Fax: +1.206.667.7264; retzioni@fhcrc.org.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

overdiagnosis based on US incidence trends, and (4) quantify the impact of control arm screening on PLCO mortality results.

Results—Screening plausibly explains 45% and changes in primary treatment can explain 33% of the US prostate cancer mortality decline. Extrapolating the ERSPC results to the long-term US setting implies an absolute mortality reduction at least 5 times greater than that observed in the trial. Approximately 28% screen-detected cases are overdiagnosed in the US versus 58% of screen-detected cases suggested by the ERSPC results. Control arm screening can explain the null result in the PLCO trial.

Conclusions—Modeling studies indicate that population trends and trial results extended to the long-term population setting are consistent with greater benefit of PSA screening—and more favorable harm-benefit tradeoffs—than has been suggested by empirical trial evidence.

Keywords

Mass screening; policy development; prostatic neoplasms; simulation modeling

Introduction

Recently, the US Preventive Services Task Force recommended against prostate-specific antigen (PSA) screening for prostate cancer based on moderate certainty that the benefits of screening do not outweigh the harms (1). The evidence of screening benefit was based primarily on mortality results from the European Randomized Study of Screening for Prostate Cancer (ERSPC) (2, 3) and the US-based Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer screening trial (4, 5). Restating the trial mortality results after medians of 13 and 11 years of follow-up respectively, the Task Force concluded that “there is adequate evidence that the benefit of PSA screening and early treatment ranges from 0 to 1 prostate cancer deaths avoided per 1,000 men screened” (1).

Although randomized trials are the gold standard for evidence-based decision making in medicine, they can have important limitations as a basis for screening policies. First, screening policy development demands information about long-term benefits and harms because these policies generally pertain to interventions conducted over an individual’s healthy lifetime. Unfortunately, most screening trials provide short-term results rather than the long-term outcomes generated by a typical population-based screening program. Second, screening trial results can be highly influenced by the trial population and by patterns of compliance with the trial protocol. Third, any inferences about screening benefit are limited to the screening strategy or strategies tested in the trial. This does not permit policy makers to identify and compare alternative policies that might be more acceptable.

In this commentary, we argue that the ERSPC and PLCO are subject to these limitations and that taking their results at face value misrepresents the likely long-term population impact of PSA screening (relative to no screening) in the US. To examine these limitations, we review modeling studies that analyze population trends and extrapolate trial evidence. A recent publication defines the models used in these studies as “mathematical frameworks that facilitate estimation of the consequences of health care decisions.” These models enable in-depth analysis of observed data that can reveal information about the disease process and facilitate extrapolation beyond the trial setting and horizon.

We summarize three main points from these studies. First, changes in primary treatments explain only a minority of the observed decline in prostate cancer mortality; assuming screening benefits would-be metastatic cases detected while still in a localized stage, PSA screening plausibly explains nearly half of the decline. Second, estimates of overdiagnoses

and lives saved based on ERSPC results overstate the likely long-term harm-benefit tradeoff in the US. Third, results from the PLCO trial, which found no significant mortality difference between the intervention and control arms, do not rule out a clinically significant benefit of PSA screening; instead, the key lesson of the PLCO trial is that more intensive screening does not necessarily yield a greater benefit than less intensive screening. This observation opens the door to investigations of more efficient screening strategies that preserve benefit while reducing cost and harms, and we can use modeling to identify such strategies.

Our discussion draws on several models developed as part of the Cancer Intervention and Surveillance Modeling Network . To model population mortality trends, we review a microsimulation model of prostate cancer progression, detection, and survival that projects mortality with and without screening or changes in treatment . We examine the same microsimulation model as well as an analytic model to estimate overdiagnosis associated with PSA screening from population incidence trends . To extrapolate long-term harm-benefit tradeoffs, we combine the estimated number overdiagnosed with the estimated lives saved due to screening, based on applying the relative reduction in prostate cancer deaths from the ERSPC, to long-term US prostate cancer mortality without screening . Finally, to interrogate PLCO results, we review a microsimulation model to replicate the PLCO trial with and without control arm contamination under a clinically significant screening benefit .

We conclude that model-based analysis of evidence is consistent with an important long-term benefit of PSA screening and that policies should be developed that improve the targeting of screening and treatment rather than eliminating the opportunity for early detection and intervention. While we focus on the case of prostate cancer screening, our arguments have bearing on the methods by which screening policies are developed, particularly when the only available trial evidence has clear limitations and extrapolation beyond the specific trial settings, horizons, and protocols is needed.

Population Outcomes

From the time of rapid adoption of PSA screening in the US in the early 1990s, prostate cancer death rates have fallen each year. The latest figures released by the Surveillance, Epidemiology, and End Results program for 2009 indicate that age-adjusted prostate cancer death rates have declined by 44% since their peak in 1991 (6). However, other aspects of prostate cancer management have also changed both before and during the PSA era. Primary treatment patterns have evolved considerably, beginning with a marked increase in the frequency of radical prostatectomy in the 1980s (7), followed by the dissemination of hormonal therapies as adjuvant to radiation therapy in the 1990s (8). Both of these treatments have been shown to be efficacious in randomized treatment trials (9, 10), although recent results from a US trial of radical prostatectomy have suggested that, in a population undergoing screening, benefit may be limited to men with intermediate or high-risk disease . In addition, radiation therapy technologies have evolved and are now able to deliver more intense, targeted doses. Advances in primary treatment have been cited as the most likely explanation for the decline in prostate cancer deaths in the US. Nonetheless, three independently developed simulation models of prostate cancer progression, diagnosis, and survival (11), making favorable assumptions regarding treatment efficacy based on the results of randomized trials and recent comparative effectiveness studies (12, 13), estimated that primary treatment changes explained only up to one-third of the drop in disease-specific deaths (11) by 2005 , leaving a large fraction of the decline to be explained by other factors.

One of the three models (14) was used to study the role of screening in the absence of treatment changes using data through 2000 and found that screening explained 45% of the

mortality decline. The model was rigorously calibrated to US prostate cancer incidence trends before and after the introduction of screening (14) and assumed that cases detected in an earlier (localized) stage by screening received a corresponding post-lead-time survival improvement. This screening benefit assumption was supported by a later analysis in which the same model was used to simulate the ERSPC and this assumption was found to be consistent with the 21% drop in prostate cancer deaths observed after a median follow-up of 11 years (15).

While these results do not prove that screening explains the observed drop in prostate cancer deaths, a decline of 70% in the incidence of advanced stage disease since 1990 supports a role for PSA screening in explaining the part of the mortality decline that is not accounted for by changes in primary treatment. A direct connection between PSA screening and the incidence of metastatic disease is evident from the ERSPC, in which the incidence of metastatic disease in the screened group (278 cases) was about half that in the control group (567 cases). While there may be other explanations for the drop in US prostate cancer deaths, these are not clearly supported by available data. For example, there have been improvements in treatment for metastatic disease, but these have been relatively recent, and a study of trends in survival among newly-diagnosed metastatic cases shows little change over time through 2005 (16). Earlier detection of progressive disease via PSA monitoring could also be a factor, but to date there have been no conclusive studies of the efficacy of earlier versus later treatment of metastatic disease following diagnosis. No etiologic factors that might have induced a change in the underlying risk of developing or dying of the disease in the population have been identified. It has been suggested that misattribution of other-cause deaths as due to prostate cancer could have caused a spurious rise in deaths due to the disease just prior to the peak in 1991, but studies of misattribution have indicated that it is a relatively rare event. We conclude that, although population prostate cancer trends are likely a complex result of many factors, modeling can help to disentangle the contributions of specific factors like treatment and screening. The observed decline in disease-specific deaths is not easily explained by other changes in disease management and is indicative of screening benefit in the population.

Trial Results I: The ERSPC

The ERSPC observed a short-term relative reduction in deaths (fraction of disease-specific deaths prevented by screening) of 21% after a median follow-up of 11 years (2, 3) and a corresponding absolute reduction (lives saved per 1,000 men enrolled in screening) of 1.07 deaths per 1,000 men invited. It is this measure of absolute benefit that formed the basis for the Task Force's conclusion that the reduction in prostate cancer mortality 10 to 14 years after PSA screening is "at most, very small" (1).

How can such a large relative effect of screening translate into such a modest absolute benefit and, ultimately, into an impression of an ineffective screening test? This is where the issues of follow-up duration (time from start of enrollment to analysis) and trial population (population from which trial participants are sampled) become essential to understand.

The absolute benefit of screening depends on the baseline disease-specific mortality rate in the unscreened population. This rate was extremely low in the ERSPC control group, amounting to only 0.5 deaths per 1,000 person years after a median 11 years of follow-up (3). Even though the 21% relative reduction was substantial, the absolute difference in rates amounted to only 0.1 deaths per 1,000 person years and this translated into 1.07 lives saved per 1,000 men screened (3).

However, in a population screening program, where men begin testing at a specified starting age and continue screening until they either die of other causes or reach an age beyond which screening is no longer recommended, the relevant baseline disease-specific mortality (17)(18)(3) approaches the lifetime probability of prostate cancer death. In addition, this lifetime probability should apply to the population of interest. The ERSPC results not only reflect short follow-up; they also pertain to a European population. In the US, in the absence of screening, the lifetime risk of a prostate cancer death based on 1990 death rates, just prior to the spread of PSA screening, was 32 per 1,000 men. A reduction of 21% yields 6.7 ($=32 \times 0.21$) lives saved per 1,000 men screened and an NNS of 149 ($=1000/6.7$). Recognizing that advances in treatment may have mitigated prostate cancer mortality, and using more recent mortality rates from 2006, we obtain a lifetime risk of prostate cancer death of 28 per 1,000 men, 5.9 lives saved per 1,000 men screened and an NNS of 170. Using endpoints of the 95% confidence interval for ERSPC mortality reduction implies a corresponding range of 2.5 ($=28 \times 0.09$) to 9.0 ($=28 \times 0.32$) lives saved per 1,000 men screened and a range of NNS of 112 to 397. These NNS estimates are considerably lower than the short-term estimates from the trial (1,410 and 1,055 at 9 and 11 years respectively).

The short-term perspective also distorts estimates of screening harm, particularly overdiagnosis. This is the detection by screening of cancers that would never have become clinically apparent without the screening test. Because prostate cancer is known to be a disease that is highly prevalent, particularly in older men, the potential for overdiagnosis is great. Indeed, after a median follow-up of 9 years, the frequency of overdiagnosis among men invited to screening was estimated to be 34 per 1,000 men screened, which was the observed excess incidence in the screened group relative to the control group (2). Combining this estimated frequency of overdiagnosis with the 9-year estimate of 0.7 lives saved per 1,000 men screened yielded a projected ratio of 48 overdiagnoses per life saved, or 48 additional cancers needed to detect (NND) to save one life (2). This figure has been cited in the media as an illustration of the high harm-to-benefit tradeoff associated with PSA screening. However, the excess incidence over the short term produces an inflated assessment of overdiagnosis amounting in this case to 58% of screen-detected cases (more than one case overdiagnosed for every two screen-detected cancers). After 11 years, the NND was revised down to 37, but both overdiagnosis and the corresponding NND are highly dependent on the setting and length of follow-up(3).

Two different models were used to estimate the long-term fraction of screen-detected cases overdiagnosed in the US population (19, 20). In contrast to the trial-based estimates, which use excess incidence in the screened group as a proxy for overdiagnosis, both models estimated the lead time associated with PSA screening (19, 20) then derived the fraction overdiagnosed as the fraction of screen-detected cases dying of other causes within their lead time. The lead-time-based estimates, namely, 23% and 28% overdiagnosed among screen-detected cases, are lower than the ERSPC estimate because they represent true overdiagnosis rather than excess incidence and because they were estimated in the US setting. Given that 16% of men will be diagnosed under current screening practices (17), then even assuming that all new cases are screen-detected and using the higher estimate of overdiagnosis, we project that 44.8 ($=0.16 \times 0.28 \times 1000$) per 1,000 will be overdiagnosed during their lifetimes. With 5.9 lives saved per 1,000 over the long-term based on US population mortality, the long-term NND is more accurately estimated as 7.6 ($=44.8/5.9$) rather than 37, giving a very different picture of the harm-benefit profile of PSA screening.

Trial Results II: The PLCO

The PLCO trial was originally designed to compare annual screening with no screening for prostate cancer. However, the trial was initiated in 1993, at the tail of the first wave of PSA

screening in the US population (21). As a consequence, screening was common among control arm subjects; an investigation into control arm contamination revealed “the intensity of PSA screening in the control arm was estimated to be approximately half of that in the intervention arm” (5). In fact, trial investigators estimated that 74% of men in the control arm received at least one routine PSA test during the 6 years of the trial compared with 95% of the intervention arm, with about half of the men in the control tested each year. Moreover, the incidence of prostate cancer in the control group was 20% higher than expected based on comparable population rates, suggesting that control arm subjects were screened even more intensively than the general population (22).

The PLCO trial has been interpreted as providing evidence that that screening is not beneficial. Indeed, the Task Force’s estimate of the mortality reduction associated with screening includes the possibility of zero lives saved over a 10- to 14-year period based on the trial result of no reduction in prostate cancer deaths in the intervention group relative to the control group. In fact, both published reports from the trial (4, 5) have shown a slight excess of prostate cancer deaths in the intervention group. When modeling was used to replicate the trial, it was found that the trial has extremely low power (range 9–25% across three models at 13 years) to infer a difference between the control and intervention groups even under a clinically significant benefit of screening (15). This is due not only to the control arm contamination but also to the lower-than-expected frequency of prostate cancer deaths in the trial population (23). Given the low numbers of deaths, the mortality results are highly variable, and the chance of excess mortality in the intervention arm, even in the presence of a significant screening benefit, is 15–29% after 13 years (15).

We conclude that the PLCO does not provide actionable information regarding screening benefit or lack thereof and that inclusion by the Task Force of zero as a potential lower bound for the benefit of screening on the basis of the reported PLCO trial results is not warranted. Instead, the study provides important evidence of the equivalence of more intensive (annual) versus less intensive (approximately biennial) screening.

Conclusions

Clinical trials are generally perceived to reflect the highest level of evidence for policy development since they are designed to eliminate systematic differences between the treatment and control groups that may lead to biased results. There have been many non-trial-based studies of PSA screening efficacy, but these have either been inconclusive or subject to extensive criticism due to their observational nature.

We are in agreement that published results from clinical trials should be used to inform policy so long as they are directly applicable to the policy-relevant setting. In the case of prostate screening and the Task Force recommendations, this setting is the US population over the long term. We have made the case that the published screening trial results do not accurately reflect the outcomes of most interest in this setting. Thus, our primary concern relates to relevance of the published trial results, and our goal is in essence to make these results more relevant for policy.

In doing so, we consider the published results to be reliable; we do not address issues of data quality although we are aware that some have questioned the quality and reporting of the ERSPC findings.

There is a further handicap associated with limiting policy development to published trials: inferences about screening harm and benefit will pertain only to those strategies actually tested. If there are strategies with more favorable harm-benefit tradeoffs than those tested, these cannot be considered or evaluated. This is critically important in the case of prostate

cancer screening with its high price tag in terms of unnecessary biopsies, which can cause serious infections including sepsis in 0.5–1.0% of biopsy cases, and overtreatment, which carries the burden of impotence and incontinence in 20–30% of cases .

There have been many suggestions for targeted screening strategies, including increasing the interval between screens (24), using more conservative criteria for biopsy referral in older men who are at higher risk of false-positive tests and overdiagnosis (25), and screening adaptively, such as lengthening the interval to the next screen if the current PSA level is below a specified threshold (26). Testing all competing policies in randomized trials is infeasible; consequently, this is a setting in which well-calibrated and validated models are proving to be of great value . We therefore advocate the use of models alongside clinical trial results, not only to interpret and extrapolate trial findings to the relevant population setting, but also to ensure that a recommendation against screening takes into account feasible alternative strategies.

(27)(27-29)The universe of models is massive and heterogeneous in terms of both model quality and methodology. Models have many limitations. They invariably represent a simplified version of disease progression, intervention, and impact. The accuracy of model results is dependent on the underlying assumptions of the model. Results themselves are subject to uncertainty which can be difficult to quantify. A recently published series of reports from a joint task force of the International Society for Pharmaceutical Outcomes Research and the Society for Medical Decision Making codifies good modeling practices and should help to insure that these limitations are acknowledged and addressed to the extent possible. However, the question of how to weigh the evidence from different data sources and study designs, including models, when formulating screening recommendations remains. Task Force recommendations for breast and colorectal screening relied on both trials and modeling studies, yet a systematic framework for synthesizing evidence from multiple study designs has yet to be established.

In conclusion, we submit these comments out of a sense that a critically important policy decision may have been made on the basis of an incomplete picture of the benefits and harms of PSA screening. We consider this to reflect a weakness not of the specific decision—a more complete picture may have led the panel to the same recommendation. Rather our concern is about the process—a process in which the most reliable evidence is defined as the observed outcome of clinical trials and there is no formal mandate to go beyond published trial estimates of benefit and harm. With a disease whose hallmark is a lengthy natural history, the harms of developing cancer screening policies based primarily on limited-duration screening trials may well outweigh the benefits.

Acknowledgments

Financial support: This work was supported by Award Numbers R01 CA131874 and U01 CA88160 from the National Cancer Institute and Award Number U01 CA157224 from the National Cancer Institute and the Centers for Disease Control. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute, the National Institutes of Health, or the Centers for Disease Control.

Funding source: This research was supported by Award Number U01 CA157224 from the National Cancer Institute and the Centers for Disease Control. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute, the National Institutes of Health, or the Centers for Disease Control. As corresponding author, Dr. Etzioni had full access to all data in the study and had final responsibility for the decision to submit for publication.

References

1. Moyer VA, on behalf of the USPSTF. Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2012
2. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med.* 2009; 360:1320–1328. [PubMed: 19297566]
3. Schroder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med.* 2012; 366:981–990. [PubMed: 22417251]
4. Andriole GL, Crawford ED, Grubb RL 3rd, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med.* 2009; 360:1310–1319. [PubMed: 19297565]
5. Andriole GL, Crawford ED, Grubb RL 3rd, et al. Prostate Cancer Screening in the Randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: Mortality Results after 13 Years of Follow-up. *J Natl Cancer Inst.* 2012; 104:1–8.
6. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2011 Sub, Vintage 2009 Pops (1973-2009) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2010 Counties. National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch; www.seer.cancer.gov released April 2012, based on the November 2011 submission
7. Lu-Yao GL, Greenberg ER. Changes in prostate cancer incidence and treatment in USA. *Lancet.* 1994; 343:251–254. [PubMed: 7905093]
8. Cooperberg MR, Grossfeld GD, Lubeck DP, et al. National practice patterns and time trends in androgen ablation for localized prostate cancer. *J Natl Cancer Inst.* 2003; 95:981–989. [PubMed: 12837834]
9. Bolla M, Gonzalez D, Warde P, et al. Improved survival in patients with locally advanced prostate cancer treated with radiotherapy and goserelin. *N Engl J Med.* 1997; 337:295–300. [PubMed: 9233866]
10. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med.* 2011; 364:1708–1717. [PubMed: 21542742]
11. Etzioni R, Gulati R, Tsodikov A, et al. The prostate cancer conundrum revisited: Treatment changes and prostate cancer mortality declines. *Cancer.* 2012
12. Cooperberg MR, Vickers AJ, Broering JM, et al. Comparative risk-adjusted mortality outcomes after primary surgery, radiotherapy, or androgen-deprivation therapy for localized prostate cancer. *Cancer.* 2010; 116:5226–5234. [PubMed: 20690197]
13. Boorjian SA, Karnes RJ, Viterbo R, et al. Long-term survival after radical prostatectomy versus external-beam radiotherapy for patients with high-risk prostate cancer. *Cancer.* 2011; 117:2883–2891. [PubMed: 21692049]
14. 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; 11:707–719. [PubMed: 20530126]
15. Gulati R, Tsodikov A, Wever EM, et al. The impact of PLCO control arm contamination on perceived PSA screening efficacy. *Cancer Causes Control.* 2012
16. Cetin K, Beebe-Dimmer JL, Fryzek JP, et al. Recent time trends in the epidemiology of stage IV prostate cancer in the United States: analysis of data from the Surveillance, Epidemiology, and End Results Program. *Urology.* 2010; 75:1396–1404. [PubMed: 19969335]
17. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Mortality - All COD, Public-Use With State, Total U.S. (1969-2005). National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch; www.seer.cancer.gov released April 2008. Underlying mortality data provided by NCHS (www.cdc.gov/nchs)
18. Hanley JA. Mortality reductions produced by sustained prostate cancer screening have been underestimated. *J Med Screen.* 2010; 17:147–151. [PubMed: 20956725]
19. Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. *Biometrics.* 2008; 64:10–19. [PubMed: 17501937]

20. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst.* 2009; 101:374–383. [PubMed: 19276453]
21. Mariotto A, Etzioni R, Krapcho M, et al. Reconstructing prostate-specific antigen (PSA) testing patterns among black and white men in the US from Medicare claims and the National Health Interview Survey. *Cancer.* 2007; 109:1877–1886. [PubMed: 17372918]
22. Pinsky PF, Black A, Kramer BS, et al. Assessing contamination and compliance in the prostate component of the Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer screening trial. *Clin Trials.* 2010; 7:303–311. [PubMed: 20571134]
23. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials.* 2000; 21:273S–309S. [PubMed: 11189684]
24. Wolf AM, Wender RC, Etzioni RB, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. *CA Cancer J Clin.* 2010; 60:70–98. [PubMed: 20200110]
25. Polascik TJ, Oesterling JE, Partin AW. Prostate specific antigen: a decade of discovery--what we have learned and where we are going. *J Urol.* 1999; 162:293–306. [PubMed: 10411025]
26. Vickers AJ, Lilja H. Urological cancer: Time for another rethink on prostate cancer screening. *Nature reviews Clinical oncology.* 2012; 9:7–8.
27. [Accessed 6/5/2012] ISPOR-SMDM Joint Modeling Good Research Practices Task Force. 2012. Available at:<http://www.ispor.org/TaskForces/GRPModelingTf.asp>
28. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2008; 149:659–669. [PubMed: 18838717]
29. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med.* 2009; 151:738–747. [PubMed: 19920274]