

Long-term projections of the harm-benefit trade-off in prostate cancer screening are more favorable than previous short-term estimates

Roman Gulati^a, Angela B. Mariotto^b, Shu Chen^a, John L. Gore^c, Ruth Etzioni^{a,*}

^aDivision of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

^bDivision of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA

^cDepartment of Urology, University of Washington, Seattle, WA, USA

Accepted 9 June 2011

Abstract

Objective: To project long-term estimates of the number needed to screen (NNS) and the additional number needed to treat (NNT) to prevent one prostate cancer death with prostate-specific antigen (PSA) screening in Europe and in the United States.

Study Design and Setting: A mathematical model of disease-specific deaths in screened and unscreened men given information on overdiagnosis, disease-specific survival in the absence of screening, screening efficacy, and other-cause mortality is presented. A simulation framework is used to incorporate competing causes of death.

Results: Assuming overdiagnosis and screening efficacy consistent with European Randomized study of Screening for Prostate Cancer (ERSPC) results, we project that, after 25 years, 262 men need to be screened and nine additional men need to be screen detected to prevent one prostate cancer death. Corresponding estimates of the NNS and the additional NNT under a range of overdiagnosis rates that are consistent with U.S. incidence are 186–220 and 2–5.

Conclusions: Long-term estimates of the NNS and the additional NNT are an order of magnitude lower than the short-term estimates published with the results of the ERSPC trial and may be consistent with cost-effective PSA screening in the general U.S. population. © 2011 Elsevier Inc. All rights reserved.

Keywords: Cost-benefit analysis; Mass screening; Prostate-specific antigen; Prostatic neoplasms; Treatment efficacy; Simulation modeling

1. Introduction

The diagnosis and treatment of prostate cancer that would never have been detected without screening—known as overdiagnosis and overtreatment—have long been recognized as the primary harms associated with prostate-specific antigen (PSA) screening. In contrast with many other cancers, overdiagnosis of prostate cancer is relatively widespread because of the high prevalence of pathological prostate cancer relative to its clinical incidence. The lifetime probability of histological prostate cancer has been estimated to be about 36% [1] whereas the lifetime probability of being diagnosed just before the advent of PSA screening was 9% [2]. These numbers suggest that

as many as three out of four prostate cancers would never become clinically apparent.

Estimates of overdiagnosis [3–7] have varied widely, primarily because of different populations and methods used for estimation. Populations may differ in their disease incidence before the introduction of screening, the screening protocol, and the intensity of diagnostic follow-up after a test. Estimating the overdiagnosis frequency is challenging because it requires making inferences about a counterfactual quantity, namely the time at which a screen-detected tumor would have been diagnosed had screening not occurred. Analytic methods typically model this latent event, but there is a range of models and, correspondingly, a range of results.

In a collaboration involving three independently developed statistical models [4,5,8], Draisma et al. [7] estimated the overdiagnosis frequency among U.S. men aged 50–84 in 1985–2000 under common assumptions about PSA utilization [9]. The models projected that 23–42% of screen-detected cases were overdiagnosed. However, because not all prostate cancer cases receive curative treatment, the

* Corresponding author. Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-B230, PO Box 19024, Seattle, WA 98109-1024, USA. Tel.: +206-667-6561; fax: +206-667-7264.

E-mail address: retzioni@fhcrc.org (R. Etzioni).

What is new?

- Early estimates of the harm-benefit trade-off from prostate cancer screening trials are based on short-term follow-up.
- These estimates, presented in terms of the number needed to screen and the additional number needed to treat to prevent one prostate cancer death, understate benefits and overstate harms.
- We use modeling to project long-term estimates of the harm-benefit trade-off from screening trial results.
- The long-term harm-benefit trade-offs of prostate cancer screening are substantially more favorable than early trial results suggested.

estimates do not directly correspond to overtreatment. In fact, data from the Surveillance, Epidemiology, and End Results (SEER) registry suggest that, if overdiagnosed cases elect primary treatment with frequencies similar to those in the general case population, then around 70% of overdiagnosed cases are overtreated, representing 16–29% of screen detections in the United States.

Estimates of overdiagnosis in the European setting are considerably higher than those in the United States. For example, Draisma et al. [4] modeled disease incidence data from the Rotterdam section of the European Randomized study of Screening for Prostate Cancer (ERSPC) and estimated that 66% of screen-detected tumors was overdiagnosed. This result is consistent with a lower baseline incidence of prostate cancer in Europe, a lower PSA cutoff for biopsy referral (3 ng/mL in several ERSPC centers vs. 4 ng/mL in the United States), and a much higher frequency of compliance with biopsy referral in the ERSPC [10] vs. the United States [11].

Although knowing the frequency of overdiagnosis is important, policy decisions about PSA screening will ultimately depend on the trade-off between harms and benefits. In 2009, estimates of screening benefit became available with the publication of interim results from the ERSPC and the U.S.-based Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer screening trial [10,12]. After a median of 9 years of follow-up, the ERSPC showed a relative risk reduction of 20% in the frequency of prostate cancer death because of screening and an absolute risk reduction of 0.71 deaths per 1,000 men screened or, reciprocally, a number needed to screen (NNS) of 1,410 to prevent one prostate cancer death. In an extended study of the Göteborg center of the ERSPC [13], PSA screening was associated with a 44% reduction in the risk of prostate cancer death after 14 years and an NNS of 293. In contrast, after 7 years of follow-up, the PLCO trial showed no difference in the prostate cancer death rates in

the screened and control arms. However, the PLCO trial was conducted in a population already undergoing screening and thus “usual care” in the control arm included screening for a substantial fraction of participants [14].

The publication that reported the ERSPC results [10] also reported a new measure of the harm-benefit trade-off associated with PSA screening, namely the “additional number needed to treat to prevent one prostate cancer death.” By comparing the excess prostate cancer incidence in the screened arm with the excess prostate cancer mortality in the control arm, they concluded that 48 additional cases of prostate cancer would potentially need to be treated to prevent one death. A similar calculation by investigators reporting the results of the Göteborg trial [13] yielded a far lower estimate of 12. Although this measure has been termed the additional number needed to treat (NNT) to prevent one death, it should, strictly speaking, be interpreted as the additional number needed to diagnose (NND) to prevent one death because presumably not all detected cases in these trials received curative treatment.

The variability in published estimates of the NNS and the additional NNT makes it difficult to determine what the harm-benefit trade-off of PSA screening is likely to be for a given setting. A major problem with these measures is their dependence on the duration of follow-up [15]. The denominator of both these measures—the absolute mortality reduction—is inevitably understated under limited follow-up. The numerator is the empirically observed excess incidence in the screened group relative to the control group, a proxy for the incidence of overdiagnosis. But this also is inaccurate under limited follow-up because excess incidence consists of a combination of overdiagnoses and early detections—cases who would have surfaced clinically in the absence of screening given sufficient follow-up time. Excess incidence over a limited time interval underestimates early detections and overestimates overdiagnosis.

In this article, we estimate the long-term NNS and additional NNT based on the mortality reduction observed in the ERSPC trial. We examine the impact of varying estimates of screening efficacy by considering efficacy results from both the ERSPC and the Göteborg trials. We use the European efficacy results because contamination in the PLCO trial prevents direct comparison of screening vs. no screening in the U.S. setting [14]. However, in an effort to examine how the NNS and additional NNT might differ in the U.S. setting, we also estimate these measures given the screening efficacy reported in the ERSPC trial but under overdiagnosis frequencies consistent with U.S. data.

2. Methods

2.1. Projecting long-term NNS and additional NNT

Long-term measures of the NNS and the additional NNT are based on long-term disease-specific mortality in the

presence and absence of screening. To project disease-specific mortality in the absence of screening, we estimate the incidence of non-overdiagnosed cancers in the trial population and model their survival based on the long-term disease-specific survival distribution without screening. To project disease-specific mortality in the presence of screening we then inflate the disease-specific survival distribution by a hazard ratio that reflects screening efficacy. The difference between the long-term, cumulative incidence of disease-specific mortality in the absence and presence of screening provides the denominator of the NNS and the additional NNT.

For the numerators, we do not yet have an estimate of overdiagnosis from the ERSPC as a whole. Draisma et al. [4] estimated the frequency of overdiagnosis in the Rotterdam section of the trial and found it to be as high as 66% of screen-detected cases. However, the excess incidence in the trial as a whole (34 per 1,000 men screened) amounted to only 58% of the screen-detected cases. This is almost certainly greater than the frequency of overdiagnosis in the trial after a median of 9 years of follow-up. In our calculations, we use this value in the absence of a more accurate estimate of overdiagnosis in the European trial. In the U.S. setting we use 23% and 42%, which correspond to the lower and higher end points of a recently published range of overdiagnosis frequencies among screen-detected cases in the United States [7].

2.2. Long-term NNS and additional NNT in the absence of other-cause mortality

We estimate the incidence of non-overdiagnosed cancer in the screened arm as:

$$N = S(1 - O) + C,$$

where O denotes the overdiagnosis frequency among screen detections and S and C denote the fraction of men screen- and non-screen-detected. For example, setting $O = 58\%$ based on Draisma et al. [4] and using $S = 4,235/72,890 \approx 5.8\%$ and $C = 1,755/72,890 \approx 2.4\%$ based on Table 1 in Schröder et al. [10], we obtain $N \approx 4.8\%$.

Because only non-overdiagnosed cases are at risk of disease-specific death, we project the long-term disease-specific survival for men on the screened arm, had they not been screened, as:

$$P = NR + (1 - N),$$

where R is the long-term disease-specific survival in the absence of screening. We use $R = 41.5\%$, the 25-year net (i.e., in the absence of other-cause death) relative survival observed for cases diagnosed in SEER in 1981–1985, just before the dissemination of PSA screening in the United States [16]. A comparison of 5-year relative survival rates between SEER 1981–1985 diagnoses and 1995–1999 diagnoses in Europe [17] yields fairly comparable results (75.8% in SEER vs. 76% in Europe). In the ERSPC

example, the resulting 25-year survival in the absence of screening is $P = 0.972$.

Next we apply the observed rate ratio H for prostate cancer death in the screened vs. the control arm to obtain long-term survival in the presence of screening. The observed mortality rate ratio is based on the entire trial population (cases and non-cases) and approximates the hazard ratio of disease-specific death associated with screening [18]. The difference in long-term survival in the presence and absence of screening becomes the denominator in the NNS and the additional NNT. Using $H = 0.80$ from the ERSPC, we estimate that the 25-year survival in the presence of screening would be $P^H = 0.972^{0.80} \approx 0.977$ (actual calculations performed using 10 significant digits), implying that the NNS is $1/(P^H - P) = 1/(0.977 - 0.972) \approx 178$ and the additional NNT is $O \times S/(P^H - P) = 0.58 \times 0.058/(0.977 - 0.972) \approx 6$. These estimates are an order of magnitude smaller than the measures reported based on interim ERSPC results. The discrepancy is partly because we are projecting survival far beyond the limited follow-up of the trial and partly because this approximation is in the absence of other-cause deaths.

2.3. Long-term NNS and additional NNT in the presence of other-cause mortality

We use a simulation-based competing risks model to incorporate other-cause deaths.

First, we simulate disease-specific and other-cause deaths without screening for a cohort representing the screened arm of the ERSPC. To do this, we (1) generate records for this cohort and randomly designate a fraction N to be non-overdiagnosed cases, (2) generate a time from diagnosis to disease-specific death for each non-overdiagnosed case using relative survival curves from SEER for cases diagnosed in 1981–1985, and (3) use life tables for men aged 55–69 at diagnosis to generate a time from diagnosis to other-cause death for each member of the cohort. Men for whom disease-specific death precedes other-cause death are assigned prostate cancer as their cause of death. We use life tables for men aged 55–69 at diagnosis to match the cohort used to derive the mortality reduction reported in Schröder et al. [10].

We use the simulated dates of disease-specific and other-cause death to construct the corresponding cumulative incidence curve for the event of prostate cancer death. Cumulative incidence of prostate cancer death stabilizes after approximately 25 years, and we use this value as our estimate of cumulative disease-specific mortality in the absence of screening.

We repeat this exercise, raising the disease-specific survival curve for the cohort to the power H to generate times to disease-specific death given screening, and we generate the corresponding cumulative incidence curve for prostate cancer death in the presence of other-cause mortality. The NNS is the reciprocal of the difference in the cumulative

Table 1. Projected 25-year NNS and additional NNT under different assumptions about the overdiagnosis rate among screen detections (*O*) and hazard ratios (*H*) for disease-specific mortality associated with screening

Projections in the absence of competing cause of death						
Overdiagnosis rate among screen detections	Hazard ratio for disease-specific mortality associated with screening	Incidence rate of non-overdiagnosed cancers in the screened arm	Disease-specific survival in screened arm in the absence of screening	Disease-specific survival in screened arm in the presence of screening	Reciprocal of deaths prevented by screening (NNS)	Overdiagnoses relative to deaths prevented by screening (additional NNT)
O	H	$N = S(1 - O) + C$	$P = NR + (1 - N)$	P^H	$1/(P^H - P)$	$O \times S/(P^H - P)$
0.58	0.80	0.048	0.972	0.977	178.35	6.01
0.58	0.56	0.048	0.972	0.984	80.79	2.72
0.42	0.80	0.054	0.968	0.975	159.48	4.45
0.42	0.56	0.054	0.968	0.982	72.21	2.01
0.23	0.80	0.069	0.960	0.968	126.27	1.69
0.23	0.56	0.069	0.960	0.977	57.11	0.76
Projections in the presence of competing causes of death						
Overdiagnosis rate among screen detections	Hazard ratio for disease-specific mortality associated with screening	Cumulative incidence of disease-specific mortality in screened arm	Cumulative incidence of disease-specific mortality in control arm	Reciprocal of deaths prevented by screening (NNS)	Overdiagnoses relative to deaths prevented by screening (additional NNT)	
O	H	M_S	M_C	$1/(M_C - M_S)$	$O \times S/(M_C - M_S)$	
0.58	0.80	0.015	0.019	262.49	8.85	
0.58	0.56	0.011	0.019	118.76	4.00	
0.42	0.80	0.018	0.023	219.83	5.36	
0.42	0.56	0.013	0.023	99.96	2.44	
0.23	0.80	0.022	0.027	185.95	2.48	
0.23	0.56	0.015	0.027	83.92	1.12	

Abbreviations: NNS, number needed to screen; NNT, number needed to treat.

incidence of prostate cancer death at 25 years in the presence of screening and in the absence of screening. The additional NNT is the overdiagnosis rate times the NNS.

2.4. Sensitivity analysis

One simplification in the proposed model is that we start the survival clock for non-overdiagnosed cases at the time of entry to the trial. Ideally, we would add an offset for these cases, representing the interval from entry to their date of clinical diagnosis. Because this interval varies from case to case and we do not have data on its distribution, we do not incorporate this offset in the model. Consequently, the age distribution for other-cause deaths likely reflects a younger age distribution (age at the start of the trial) than the one that is the case in practice (age at diagnosis). To assess the impact of this simplification, we redid our calculations assuming an age distribution for competing mortality that was 5 years older than the one used in the baseline model.

A related simplification is that we do not explicitly account for lead time among screen detections. To assess the impact of this simplification, we generated results under 20-year net relative survival for screen detections and 25-year net relative survival for non-screen-detected cases. By doing this, we effectively assigned screen-detected cases the long-term survival that would correspond to a 5-year lead time.

3. Results

The upper panel of Table 1 presents calculations of NNS and additional NNT in the absence of competing causes of death. We observe that a higher overdiagnosis rate implies better long-term, disease-specific survival (P) in the absence of screening. Given a specified mortality hazard ratio (H), this produces a lower number of deaths prevented, a higher NNS, and a higher additional NNT. These relationships also hold in the presence of other-cause death as shown in the lower panel, which has fewer prostate cancer deaths in both the absence and the presence of screening and correspondingly fewer prostate cancer deaths prevented [19,20].

Assuming screening efficacy as reported in the ERSPC, we find that, after 25 years, the NNS is 262 and the additional NNT is nine in the European setting. Although this is based on our highest setting for the overdiagnosis frequency, the additional NNT is much lower than the estimate of 48 presented in Schröder et al [10]. Corresponding estimates in the U.S. setting are 186–220 and 2–5. Accounting for a 5-year delay between entering the trial and diagnosis increases the additional NNT by about 12%, and accounting for a 5-year lead time increases the additional NNT by 12–15%.

4. Discussion

The NNS and the additional NNT have the potential to be highly informative about the harm-benefit trade-off of PSA screening. However, the additional NNT of 48 [10] has been repeatedly misinterpreted, often justifying an overly negative impression of the trade-off [21–23]. The additional NNT is not a measure of overdiagnosis and should not be misinterpreted as such; it is the ratio of a specific harm of screening (overdiagnosis) to a specific benefit (deaths prevented). Moreover, its value depends not only on the efficacy of screening but also on the extent of overdiagnosis and the time horizon.

By estimating the long-term NNS and additional NNT, we have eliminated variability because of the follow-up duration, allowing us to study their dependence on overdiagnosis and screening efficacy. We find that even after using the excess incidence observed in the ERSPC as an upper bound for overdiagnosis, the resulting NNS and additional NNT remain an order of magnitude lower than the published estimates from the ERSPC as a whole. Using a different method, a recent study found that these measures decreased substantially even after only 12 years of follow-up [15].

Our additional NNT estimates are also lower than those of Welch and Albertsen [24], who obtained estimates of 23 for the additional NND and 18 for the additional NNT in the United States. We note, however, that although these results are predicated on an apparently optimistic assumption, namely that the entire disease-specific mortality decline is attributable to screening, the corresponding deaths prevented are computed over an interval that ends in 2005, and that certainly excludes many deaths prevented by screening that took place toward the end of this interval. Thus, this estimate of deaths prevented may ultimately underestimate the true mortality benefit because of screening over the calendar interval considered. In addition, Welch and Albertsen use excess incidence (relative to the incidence just before the advent of PSA screening) rather than true overdiagnosis as their additional NND and additional NNT numerators. We conclude that their results, although lower than 48, likely also overestimate the true additional NND and additional NNT in the United States.

The NNT is established as a measure of harm-benefit trade-off in the setting of treatment trials, where it serves as an interpretable summary of crude (in the presence of other-cause death) survival differences among the groups being compared. Its analog in the screening trial setting is the NNS, which similarly depends only on survival in the screening and control groups. It is important to recognize that the NNS depends not only on the screening protocol implemented in practice, but also on the frequency of treatment in the trial population [25], and therefore it may not be transferable to populations that differ from the trial population in these respects. Although the NNS and the additional NNT are closely related, they are different

measures and both should be considered when evaluating the harms and benefits of a screening policy.

We conclude by noting that a recent cost-effectiveness study suggests that the projected long-term additional NNT estimated in this article for PSA screening is sufficiently low for PSA screening to satisfy conventional thresholds for cost-effectiveness in the U.S. population [26]. It is likely that even greater cost-effectiveness will be achieved under screening strategies targeted at younger and healthier men [27]. However, additional modeling will be necessary to go beyond empirical observation of screening trials with specified protocols and to determine optimal screening strategies in the general population.

Acknowledgments

The authors thank Gerrit Draisma for helpful comments on an earlier draft of this manuscript.

Sources of support: This research was supported by Award Number 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.

References

- [1] Etzioni R, Cha R, Feuer EJ, Davidov O. Asymptomatic incidence and duration of prostate cancer. *Am J Epidemiol* 1998;148:775–85.
- [2] Surveillance, Epidemiology, and End Results (SEER) Program SEER*Stat Database: Incidence - SEER 9 Regs Limited-Use, Nov 2006 Sub (1973–2005), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2008, based on the November 2007 submission; [cited]. Available at www.seer.cancer.gov. Accessed August 9, 2011.
- [3] McGregor M, Hanley JA, Boivin JF, McLean RG. Screening for prostate cancer: estimating the magnitude of overdiagnosis. *CMAJ* 1998;159:1368–72.
- [4] Draisma G, Boer R, Otto SJ, van der Cruijsen IW, Damhuis RA, Schröder FH, et al. Lead times and overdiagnosis due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003;95:868–78.
- [5] Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst* 2002;94:981–90.
- [6] Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. *Biometrics* 2008;64:10–9.
- [7] Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Gulati R, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst* 2009;101:374–83.
- [8] Tsodikov A, Szabo A, Wegelin J. A population model of prostate cancer incidence. *Stat Med* 2006;25:2846–66.
- [9] Mariotto A, Etzioni R, Krapcho M, Feuer EJ. 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;109:1877–86.
- [10] Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320–8.
- [11] Pinsky PF, Andriole GL, Kramer BS, Hayes RB, Prorok PC, Gohagan JK. Prostate biopsy following a positive screen in the Prostate, Lung, Colorectal and Ovarian cancer screening trial. *J Urol* 2005;173:746–50. discussion 50–51.
- [12] Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009;360:1310–9.
- [13] Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, et al. Mortality results from the Göteborg randomised population-based prostate-cancer screening trial. *Lancet Oncol* 2010;11:725–32.
- [14] Pinsky PF, Black A, Kramer BS, Miller A, Prorok PC, Berg C. Assessing contamination and compliance in the prostate component of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *Clin Trials* 2010;7:303–11.
- [15] Loebl S, Vonesh EF, Metter EJ, Carter HB, Gann PH, Catalona WJ. What is the true number needed to screen and treat to save a life with prostate-specific antigen testing? *J Clin Oncol* 2011;29:464–7.
- [16] Surveillance, Epidemiology, and End Results (SEER) Program SEER*Stat Database: Incidence - SEER 17 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Linked To County Attributes - Total U.S., 1969–2007 Counties [database on the Internet]. Released April 2010, based on the November 2009 submission (1973–2007 varying) [cited]. Available at www.seer.cancer.gov.
- [17] Berrino F, De Angelis R, Sant M, Rosso S, Bielska-Lasota M, Coebergh JW, et al. Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995–99: results of the EUROCARE-4 study. *Lancet Oncol* 2007;8:773–83.
- [18] Whitehead J. Fitting Cox's regression model to survival data using GLIM. *J R Stat Soc Ser C Appl Stat* 1980;29:268–75.
- [19] Cronin KA, Feuer EJ. Cumulative cause-specific mortality for cancer patients in the presence of other causes: a crude analogue of relative survival. *Stat Med* 2000;19:1729–40.
- [20] Welch HG, Albertsen PC, Nease RF, Bubolz TA, Wasson JH. Estimating treatment benefits for the elderly: the effect of competing risks. *Ann Intern Med* 1996;124:577–84.
- [21] Kolata G. Studies show prostate test saves few lives. *New York Times*. 2009 March 19.
- [22] Ablin R. The great prostate mistake. *New York Times*. 2010 March 9.
- [23] Woolston C. The debate over prostate cancer tests. *Los Angeles Times*. 2011 March 7.
- [24] Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986–2005. *J Natl Cancer Inst* 2009;101:1325–9.
- [25] Richardson A. Screening and the number needed to treat. *J Med Screen* 2001;8:125–7.
- [26] Shteynshlyuger A, Andriole GL. Cost-effectiveness of prostate specific antigen screening in the United States: extrapolating from the European study of screening for prostate cancer. *J Urol* 2011;185:828–32.
- [27] Crawford ED, Grubb R 3rd, Black A, Andriole GL Jr, Chen MH, Izmirlian G, et al. Comorbidity and mortality results from a randomized prostate cancer screening trial. *J Clin Oncol* 2011;29:355–61.