Estimating the harms and benefits of prostate cancer screening as used in common practice versus recommended good practice

A microsimulation screening analysis

Carlsson, Sigrid V.; de Carvalho, Tiago M.; Roobol, Monique J.; Hugosson, Jonas; Auvinen, Anssi; Kwiatkowski, Maciej; Villers, Arnauld; Zappa, Marco; Nelen, Vera; Páez, Alvaro; Eastham, James A.; Lilja, Hans; de Koning, Harry J.; Vickers, Andrew J.; Heijnsdijk, Eveline A M

Published in:
Cancer

DOI:
10.1002/cncr.30192

2016

Document Version:
Peer reviewed version (aka post-print)

Estimating the Harms and Benefits of Prostate Cancer Screening As Used in Common Practice Versus Recommended Good Practice: A Microsimulation Screening Analysis

Sigrid V. Carlsson, MD, PhD, MPH[1-3]; Tiago M. de Carvalho, PhD student[4]; Monique J. Roobol, PhD[5]; Jonas Hugosson, MD, PhD[3,6]; Anssi Auvinen, MD, PhD[7]; Maciej Kwiatkowski, MD[8,9]; Arnauld Villers, MD[10]; Marco Zappa, MD[11]; Vera Nelen, MD[12]; Alvaro Páez, MD[13]; James Eastham, MD[14]; Hans Lilja, MD, PhD[14-18]; Harry J. de Koning, MD, PhD[4]; Andrew J. Vickers, PhD[2]; and Eveline A.M. Heijnsdijk, PhD[4]

[1]Dept. of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY *[MSKCC]*

[2]Dept. of Epidemiology & Biostatistics, MSKCC

[3]Department of Urology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden *[Sahlgrenska Academy]*

[4]Dept. of Public Health, Erasmus Medical Center, Rotterdam, The Netherlands *[Erasmus]*

[5]Dept. of Urology, Erasmus

[6]Sahlgrenska University Hospital, Sahlgrenska Academy

[7]Tampere University, School of Health Sciences, Tampere, Finland

[8]Department of Urology, Kantonsspital Aarau, Aarau, Switzerland

[9]Department of Urology, Academic Hospital Braunschweig, Brunswick, Germany

[10]Department of Urology, CHU Lille, University of Lille Nord de France, Lille, France

[11]Unit of Clinical and Descriptive Epidemiology, Istituto per lo Studio e la Prevenzione Oncologica, Florence, Italy

[12]Provinciaal Instituut voor Hygiene, Antwerp, Belgium

[13]Department of Urology, Hospital Universitario de Fuenlabrada, Madrid, Spain

[14]Urology Service at the Department of Surgery, MSKCC

[15]Department of Laboratory Medicine, MSKCC

[16]Genitourinary Oncology Service, Dept. of Medicine, MSKCC

[17]Nuffield Dept. of Surgical Sciences, University of Oxford, Oxford, UK

[18]Dept. of Translational Medicine, Lund University, Malmo, Sweden

Running head: Estimating net benefit of PSA screening

Precis for use in TOC: To compare the benefit and harm of PSA-based prostate cancer screening, the authors used MIcrosimulation SCreening ANalysis and compared common practice to recommended "good practice." They found that common screening and treatment practices are associated with little net benefit, whereas following a few straightforward clinical recommendations, particularly greater use of active surveillance for low-risk disease and reducing screening in older men, would lead to an almost 4-fold increase in the net benefit of PSA screening.

Corresponding author: Sigrid V. Carlsson, MD, PhD, MPH, Memorial Sloan Kettering Cancer Center, Departments of Surgery and Epidemiology & Biostatistics, 485 Lexington Avenue, New York, NY 10065, USA; carlssos@mskcc.org; Phone: 1-646-888-8250

Total word count: 5398/5500

Tables: 5

Figures: 0

Authors' contributions

AJV conceived the study. AJV, SVC, JE, HL, JH, and MJR designed the recommended good practice recommendations. EH, HJK, and TMC conceived, designed and calibrated the MISCAN model. SVC, AJV, and EH performed the literature search. SVC, MJR, JH, AA, MK, AV, MZ, VN, AP, and HL are all members of the ERSPC trial, and contributed to the data acquisition upon which MISCAN builds. SVC, TMC, MJR, HJK, EH, and AJV carried out the data analyses. EH ran the MISCAN model. SVC and AJV drafted the manuscript. All authors read, interpreted, and edited the manuscript. EH had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final submitted manuscript version.

Conflict of interest

Monique J. Roobol serves on the advisory board of Opko Health. Anssi Auvinen reports personal fees from EPID Research and lecture fees from GlaxoSmithKline outside the submitted work. Maciej Kwiatkowski reports personal/consulting fees from Astellas, Janssen, and Myriad Genetics outside the submitted work. Hans Lilja reports service on a Roche Diagnostics advisory panel, outside the submitted work; he has an immediate family member employed at Ferring Pharmaceuticals; he holds patents for free PSA, hK2, and intact PSA assays (licensed and commercialized by Opko Health) and is named with Andrew J. Vickers on a patent application for a statistical method to detect prostate cancer, which has been commercialized by Opko Health (both authors receive royalties from sales of the tests); and he owns stock in Opko Health. Harry J. de Koning received support from a research grant consulting fees from Beckman Coulter paid to the Department of Public Health at Erasmus Medical Center. Andrew J. Vickers serves as a consultant to Genome DX and Genomic Health, outside the submitted work; he is named with Hans Lilja on a patent application for a statistical method to detect prostate cancer, which has been commercialized by

Opko Health (both authors receive royalties from sales of the test); and he holds Opko Health stock options. Eveline A. M. Heijnsdijk received support from a research grant from Beckman Coulter paid to the Department of Public Health at Erasmus Medical Center.

Funding: This work was supported by grants from AFA Insurance, the Swedish Cancer Society, the Swedish Prostate Cancer Foundation, the Research Foundation at the Department of Urology at Sahlgrenska University Hospital, Sweden America Foundation, the Swedish Council for Working Life and Social Research, and the Swedish Society for Medical Research (to Sigrid V. Carlsson). Sigrid V. Carlsson, James Eastham, Hans Lilja, and Andrew J. Vickers are supported in part by a Cancer Center Support Grant from the National Institutes of Health/National Cancer Institute (NIH/NCI) made to Memorial Sloan Kettering Cancer Center (P30 CA008748). Additional support was received from the Sidney Kimmel Center for Prostate and Urologic Cancers and from David H. Koch through the Prostate Cancer Foundation. The NIH/NCI supported the work with grants P50 CA092629 and R01 CA160816. This publication was made possible by Grant Number U01 CA157224 from the National Cancer Institute as part of the Cancer Intervention and Surveillance Modeling Network (CISNET), which supported a forum for the comparative development of simulation-based decision models. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute. Additional support was provided by the National Institute for Health Research Oxford Biomedical Research Centre Program; the Swedish Cancer Society (Cancerfonden project no. 14-0722); an FiDIPro-program award from TEKES, Finland; and Fundacion Federico SA; a research grant from Beckman Coulter (to Eveline A. M. Heijnsdijk and Harry J. de Koning); the Dutch Cancer Society and the Netherlands Organization for Health Research and Development (to Harry J. de Koning); and the Finnish Cancer Society and the Academy of Finland (to Anssi Auvinen).

ABSTRACT

BACKGROUND: Prostate-specific antigen (PSA) screening and concomitant treatment can be implemented in several ways. We investigated how the net benefit of PSA screening varies between common practice versus "good practice."

METHODS: We used MIcrosimulation SCreening ANalysis (MISCAN) to evaluate the effect on quality-adjusted life-years (QALYs) if 4 recommendations were followed: limited screening in older men; selective biopsy in men with elevated PSA; active surveillance for low-risk tumors; and treatment preferentially delivered at high-volume centers. Outcomes were compared to a base model with annual screening starting at ages 55–69, simulated using the European Randomized Study of Screening for Prostate Cancer (ERSPC) data.

RESULTS: In terms of QALYs gained compared to no screening, per 1000 screened men followed over their lifetime, recommended good practice led to 73 life-years (LYs) and 74 QALYs gained compared to 73 LYs and 56 QALYs for the base model. In contrast, common practice led to 78 LYs gained but only 19 QALYs gained; more than a 75% relative reduction in QALYs gained from unadjusted LYs gained. The poor outcomes for common practice were influenced predominantly by use of aggressive treatment for low-risk disease, with PSA testing in older men also strongly reducing potential QALY gains.

CONCLUSIONS: Commonly-used PSA screening and treatment practices are associated with little net benefit. Following a few straightforward clinical recommendations, particularly greater use of active surveillance for low-risk disease and reducing screening in older men, would lead to an almost 4-fold increase in the net benefit of prostate cancer screening.

Keywords (MeSH): Prostate-Specific Antigen/blood, Prostatic Neoplasms, Mass Screening, Quality-of-Life, Quality-Adjusted-Life-Years, Early Detection of Cancer/adverse effects

INTRODUCTION

The European Randomized Study of Screening for Prostate Cancer (ERSPC) demonstrated that regular prostate-specific antigen (PSA) screening every 2-4 years leads to a relative reduction in prostate cancer (PC)-specific mortality of 21% at 13 years of follow-up.[1] However, this benefit is offset by harms, in terms of over-diagnosis and consequent side-effects from treatment, hence the clear recommendation against PSA screening from the United States Preventive Services Task Force in 2012.[2] Using MIcrosimulation SCreening ANalysis (MISCAN), we have previously shown that over a lifetime, screening leads to a 28% relative reduction in PC-specific mortality and 8.4 life-years gained per averted death.[3] However, this benefit is mitigated by a loss in quality-adjusted life-years (QALYs)—a 23% reduction from life-years gained—primarily because of side-effects of treatment such as urinary and erectile dysfunction.[3]

There have been considerable advances in our understanding of PC and PSA since the ERSPC was initiated in the early 1990s. Empirical data suggest that the ratio of benefit-to-harm could be improved by restricting screening to appropriate age ranges, restricting biopsy and treatment to men at highest risk, and shifting treatment to higher-volume centers.[4-6] These relatively uncontroversial findings have been incorporated in many guidelines. In contrast, research into common clinical practice has found frequent PSA testing among older men with limited life expectancy,[7-8] aggressive use of curative treatment for low-risk tumors,[9] and surgical treatment largely performed by low-volume providers.[10]

We hypothesize that the benefit-to-harm ratio from PSA screening and subsequent treatment would be improved by following a straightforward set of simple good practice guidelines. We sought to quantify the effects of implementing these recommendations upon the outcomes of PC screening using MISCAN. We compared a “recommended good practice” model versus a model reflecting common screening and treatment practices, with a base model using ERSPC data.

METHODS

MISCAN

The MISCAN model, described in detail elsewhere,[3] simulates individual life histories with and without PSA screening, and with and without development of PC. The "tumor growth model" simulates PC natural history, which progresses from no disease, to preclinical screen-detectable PC, to clinical cancer at various stages. Thereafter, the tumor is screen-detected, clinically diagnosed, or progresses to another stage. The model is calibrated using raw data from the core age group (55–69 years) of the Rotterdam and Göteborg sections of the ERSPC. This includes follow-up data through 2008 (median 11 years) and a stage-dependent cure rate estimated for the observed PC-specific mortality reduction of 29% among attendees to screening in ERSPC.[3] The model was subsequently validated using data from all centra in the ERSPC, for both the screening and the control arms (thus accounting for a low contamination rate), as described earlier.[3]

The effectiveness of radical prostatectomy (RP) compared to watchful waiting was assigned a relative risk of PC-specific mortality of 0.65 based on Scandinavian Prostate Cancer Group-4 data[11]; a similar effect was assumed for radiotherapy (RT). Survival was modeled using the Gleason score−dependent Albertsen data[12] as well as Surveillance, Epidemiology, and End Results (SEER) data.[3]

QALYs were calculated by multiplying utility estimates for various health states, where 0 is death (or worst imaginable health) and 1 is full health, by the duration and number of men in the state. Utility estimates were obtained from the Cost-Effectiveness Analysis Registry[13] or the literature, or were based on assumptions. For active surveillance (AS), we assumed an estimated utility of 0.97 for the base case. A complete justification and references to the assumptions used in the base model were reported previously.[3]

Model building

MISCAN relies on certain parameter inputs, which can be changed. We simulated lifetime outcomes for those who underwent PSA screening versus controls who did not undergo screening, for a male population aged 0 to 100 years, with an age distribution based on the European Standard Population.[3] We changed some of MISCAN's inputs to investigate the effects of the different models upon QALYs.

The base model uses annual PSA screening, as often practiced in the U.S. It follows a population of men aged 0–100 over their lifetimes and screens them, with 80% participation rate, between ages 55 and 69; matching the ERSPC core age group where a significant effect on PC-specific mortality was demonstrated in favor of screening.[1, 14-15] The base model also uses: positive predictive value (PPV) of biopsy of 22.7% as in the ERSPC; primary treatment distribution (RP, RT, or AS with deferred treatment) based on age, T-stage and Gleason score as in the ERSPC; and complication rates after curative treatment as seen in U.S. population-based series.[3]

We created 2 additional models: "recommended good practice," which amended the base model by incorporating 4 simple recommendations on screening and treatment found in many guidelines, and "common practice," in which we incorporated data from empirical studies of contemporary U.S. practice patterns. Table 1 lists the assumptions changed from the base model.

Age for screening. The ERSPC found no evidence of benefit for men who start PSA screening at age ≥70, with the lower bound of the 95% CI excluding the central estimate for risk reduction for men aged <70.[15] Similarly, the American Urological Association does not recommend routine PSA screening in men aged ≥70 years.[16] For the common practice model, where some men were assumed to continue screening after age 70, we used age-dependent screening rates from an empirical study of health behaviors in the U.S.: ages 70–74: 47%; 75–79: 44%; 80–84: 43%; 85+:

26%.[7] As that study included all ages over 84 into a single category, we assumed the 26% rate of screening for this category applied to ages 85–90, with no screening above age 90.

Biopsy criteria. The ERSPC study protocol stated that men with a positive screening test (PSA ≥ 3.0 ng/mL) should be recommended for biopsy. The proportion of test-positive men who had evidence of cancer on biopsy was only 22.7%.[14] In common urologic clinical practice, patients with elevated PSA are evaluated for benign disease and subject to repeat PSA testing before the decision to biopsy.[17] We investigated how screening outcomes would change if men with elevated PSA were biopsied more selectively, based on clinical work-up. Instead of a PPV of 22.7% for biopsy after a positive PSA test, we applied a PPV of 40%, in line with U.S. clinical cohorts,[18] for both the "recommended good practice" model and the "common practice" model.

Active surveillance (AS). Recent data clearly indicate that not all men with PC need immediate treatment, and low-risk tumors can be safely managed by the approach known as active surveillance, with repeat biopsy and routine monitoring of the disease.[19] Several guideline groups, such as the National Comprehensive Cancer Network, now recommend AS for low-risk PC.[17] We investigated how QALYs were affected if men with low-risk disease (clinical stage T1, Gleason score 6) were enrolled in AS. In the base model, AS usage depended on age, and averaged 30% across all men with low-risk tumors. For cumulative proportions of men leaving AS each year, we used data from Klotz's series: year 1: 8%, year 2: 16%, year 3: 20%, year 4: 24%, year 5: 28%, year 6: 29%, and year 7: 30%.[19] For the recommended good practice model, we applied a 90% rather than a 100% AS rate to men with low-risk disease, given that there may be clinical reasons to treat some low-risk men. For the common practice model, we applied an AS rate of 9.2% for men with low-risk disease, obtaining this estimate from the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE) registry 1990–2008, and also reflecting what has been practice for many years.[9]

High-volume centers. There is a considerable literature on the volume-outcome relationship, suggesting decreased complications and side-effects and improved outcomes for patients treated by high-volume providers.[20-23] Shifting treatment trends so that more patients are treated by high-volume surgeons could, therefore, possibly improve cancer control and decrease complications. There have been widespread calls for "regionalization"[24]; that is, increasing the proportion of patients treated at high-volume centers.[25] We investigated how QALYs were affected if impotence and incontinence rates after RP were in line with rates seen at high-volume centers.[26] The MISCAN model used a representative, multiregional, U.S. cohort as the source for estimates of overall sexual problems and urinary leakage problems at 24 months post-RP, taking baseline functioning into account.[27] The base model assumed 30% overall sexual bother, 6% urinary bother, and 0% bowel bother post-RP.[3] Although different rates were used for RT (20% sexual, 5% urinary, and 8% bowel bother), when multiplied with utilities, total utility ended up being similar for the two treatment modalities. These estimates may seem lower than many reported in the literature because they are marginal—that is, they take into account that some men would have dysfunction without surgery/ RT. Also, these estimates reflect bother not function, and not all men experiencing dysfunction report lowered utility.

Estimates for functional outcomes after RP for surgeons at a high-volume center were derived from empirical data using case-mix-adjusted outcomes,[26], giving rates of sexual and leakage problems of 19% and 5%, respectively.

Sensitivity analysis

A sensitivity analysis was performed, comparing QALYs gained between the 3 models. In an attempt to reflect the effect of the different strategies on a population level, rather than an individual level, we varied the utility estimates (more vs less extreme) by about half those previously published.[3]

We compared 5 different scenarios per model using different combinations of utilities for screening procedures versus treatment and terminal illness, ie, reflecting varying population-level trade-offs for tolerability of screening procedures versus down-stream consequences.

Since the use of AS for men with low-risk disease has increased over the past years, another sensitivity analysis was performed, with a 34% AS rate, as reported in a recent update from the CaPSURE registry for the time period 2008–2013.[28]

RESULTS

Effect of modeling on QALYs

Table 2 shows quality-adjusted effects of the 3 screening models, compared to no screening, given various health states. The recommended good practice model displayed favorable effects at the biopsy stage. Compared to the base model, the good practice model had more QALYs lost in the AS health state due to its increased AS rate (3.2 vs. 9.7 QALYs per 1000 men); however, this was balanced by fewer QALYs lost from side-effects after RT and RP. The opposite was true for the common practice model, with few QALYs lost for AS, but substantial losses in QALYs due to the higher rate of treatment with RT and RP.

The predicted effects of the screening approaches are shown in Table 3. Compared to the base model, recommended good practice led to an improvement in QALYs gained, from 56 to 74, largely related to increased use of AS. This approach also substantially reduced the number of biopsies performed, from 605 to 407 per 1000 men. In contrast, common practice with screening up to age 90 years and with a 9.2% AS rate, led to 78 life-years gained but only 19 QALYs gained. This is more than a 75% relative reduction in QALYs gained from unadjusted life-years gained. Of the QALYs lost by following common practice compared to recommended good practice, about 24 were related to overtreatment of low-risk disease, 34 due to screening older men, and 3 due to

treatment at low-volume centers (Table 4). Note that these figures do not add up to the 55 QALY difference between common practice and recommended good practice because of interaction effects, such as the impact of overtreatment in older men.

In a sensitivity analysis varying the more and less extreme utility estimates in an attempt to reflect the effect on QALYs of the different strategies at a population level, did not show recommended good practice leading to worse outcomes than the base or common practice models (Supplementary material).

Increasing the use of AS to 34%, to reflect more contemporaneous rates, yielded an overall 30 QALYs gained for current practice compared to 74 QALYs for recommended good practice.

DISCUSSION

This study examined the effect upon QALYs of widespread implementation of 4 widely-accepted screening and treatment recommendations, compared to common clinical practice. Microsimulation modeling showed that care following the good practice recommendations – restricting screening in elderly men, selective biopsy, AS for low risk tumors and preferential referral to high-volume centers – led to a large improvement in QALYs gained per 1000 men, up to 74 from 56 for the base model. In contrast, common screening and treatment practice was estimated to lead to only 19 QALYs gained, translating into a more than 75% relative loss in potential QALYs gained.

Naturally, any modeling study is only as good as the model used. The MISCAN model has been shown to adequately predict PC incidence and PC-specific mortality in the Netherlands.[3] When applied to the U.S. population and compared to other models, differences are relatively minor (eg, lead time of 7.9 vs. 6.9 years). In comparison with 2 other models, MISCAN may be conservative, that is, may overestimate some screening harms.[29] We have also previously argued

that the European data may underestimate the benefits of screening due to sub-optimal treatment efficacy in the ERSPC, where both radiation doses and surgeon volumes were much lower than would be optimal.[30] Note that we did not include higher cure rates associated with referral to high-volume centers in our "recommended good practice" model, perhaps underestimating the benefits for more regionalized care. Furthermore, the differences in urinary and sexual problems between standard care and care at high-volume centers were relatively modest in our model: 1% and 11%, respectively, in absolute terms. Again, this may lead to some underestimation of the effects of regionalized treatment.

There has been considerable recent interest in the use of risk-stratified methods of evaluating men with elevated PSA-levels before biopsy, such as reflex blood tests or multiparametric magnetic resonance imaging. The PPV associated with these tests is likely even higher than the 40% figure used in our models. The QALYs gained with recommended good practice may, therefore, be a slight underestimation. However, we do not expect this to make a large difference to our findings as the near 20-point increase in PPV used in the main analysis led to only a minor improvement in QALYs gained (+1.2 QALYs).

There is some evidence that current practice is changing. Across community-based urology practices in Michigan, half of men with low-risk PC now receive initial AS.[31] We expect there will be more pronounced changes throughout the U.S. in the near future. Changing the use of AS to 34%, as reported in the most recent update from the CaPSURE registry,[28] did increase overall QALYs gained from 19 to 30. These are promising signs that changes in urologic practice will make a large difference to quality-of-life outcomes of screening.

There is also evidence that screening practices in older men have been changing for the better. For instance, incidence data from SEER have indicated that the age-, race- and ethnicity-adjusted rate of early-stage PC among men ≥75 fell from 443 to 330 per 100,000 (−25.4%; $p<0.001$)

between 2007 and 2009.[32] While encouraging, these changes go only a small way toward the major shift in screening and treatment practices needed for U.S. practice to be compliant with good practice recommendations.

Critics of PSA screening claim that it has little benefit and causes significant harm. This may be the case as PSA screening is currently implemented in the US, but does not take into account the potential benefit of screening that follows good practice recommendations. Addressing the problems of screening in older men and aggressive treatment of low-risk disease might be expected to strongly increase the benefit of PSA screening.

A limitation of the present study is that results based on the MISCAN model are relevant for Caucasian men and may not apply to men of other ethnicities.

CONCLUSIONS

Common practices for PSA screening and subsequent PC treatment are associated with considerable harm and moderate benefit. Changing practices to conform to established recommendations would lead to an estimated 4-fold increase in the net benefit of screening.

REFERENCES

1. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014;384:2027–35.

2. Moyer VA; U.S. Preventive Services Task Force. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2012;157:120–34.

3. Heijnsdijk EA, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. *N Engl J Med*. 2012;367:595–605.

4. Carlsson S, Vickers AJ, Roobol M, et al. Prostate cancer screening: facts, statistics, and interpretation in response to the U.S. Preventive Services Task Force Review. *J Clin Oncol*. 2012;30:2581–4.

5. Vickers A, Carlsson S, Laudone V, and Lilja H. It Ain't What You Do, It's the Way You Do It: Five Golden Rules for Transforming Prostate-Specific Antigen Screening. *Eur Urol*. 2014;66:188–90.

6. Vickers AJ, Sjoberg DD, Ulmert D, et al. Empirical estimates of prostate cancer overdiagnosis by age and prostate-specific antigen. *BMC Med*. 2014;12:26.

7. Drazer MW, Huo D, Schonberg MA, et al. Population-based patterns and predictors of prostate-specific antigen screening among older men in the United States. *J Clin Oncol*. 2011;29:1736–43.

8. Drazer MW, Prasad SM, Huo D, Razmaria A, Eggener SE. National trends in prostate cancer screening among older American men with limited 9-year life expectancies: Evidence of an increased need for shared decision making. *Cancer*. 2014;120:1491–8.

9. Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate cancer. *J Clin Oncol*. 2010;28:1117–23.

10. Savage CJ, Vickers AJ. Low annual caseloads of United States surgeons conducting radical prostatectomy. *J Urol*. 2009;182:2677–2679.

11. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med*. 2011;364:1708–17.

12. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. *JAMA*. 2005;293:2095–101.

13. Institute for Clinical Research and Health Policy Studies. The Cost-Effectiveness Analysis Registry. Boston: Tufts Medical Center, Center for the Evaluation of Value and Risk in Health. Available at: http://www.cearegistry.org

14. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360:1320–8.

15. Schröder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med*. 2012;366:981–90.

16. Carter HB, Albertsen PC, Barry MJ, et al. Early detection of prostate cancer: AUA Guideline. *J Urol.* 2013;190:419–26.

17. National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology (NCCN Guidelines) – Prostate Cancer Early Detection, version 2.2015. Available at: http://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf

18. Vickers AJ, Cronin AM, Roobol MJ, et al. The relationship between prostate-specific antigen and prostate cancer risk: the Prostate Biopsy Collaborative Group. *Clin Cancer Res*. 2010;16:4374–81.

19. Klotz L, Zhang L, Lam A, Nam R, Mamedov A, Loblaw A. Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. *J Clin Oncol*. 2010;28:126–31.

20. Begg CB, Riedel ER, Bach PB, et al. Variations in morbidity after radical prostatectomy. *N Engl J Med*. 2002;346:1138–44.

21. Vickers AJ, Bianco FJ, Serio AM, et al. The surgical learning curve for prostate cancer control after radical prostatectomy. *J Natl Cancer Inst*. 2007;99:1171–7.

22. Trinh QD, Bjartell A, Freedland SJ, et al. A systematic review of the volume-outcome relationship for radical prostatectomy. *Eur Urol*. 2013;64:786–98.

23. Eastham JA. Do high-volume hospitals and surgeons provide better care in urologic oncology? *Urol Oncol*. 2009;27:417–21.

24. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med*. 1979;301:1364–1369.

25. Milstein A, Galvin RS, Delbanco SF, Salber P, Buck CR Jr. Improving the safety of health care: The Leapfrog Initiative. *Eff Clin Pract*. 2000;6:313–316.

26. Vickers A, Savage C, Bianco F, et al. Cancer control and functional outcomes after radical prostatectomy as markers of surgical quality: analysis of heterogeneity between surgeons at a single cancer center. *Eur Urol*. 2011;59:317–22.

27. Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. *N Engl J Med*. 2008;358:1250–61.

28. Cooperberg MR, Carroll PR. Trends in Management for Patients With Localized Prostate Cancer, 1990-2013. *JAMA*. 2015;314:80–2.

29. 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;101:374–83.

30. Vickers AJ, Lilja H. Prostate cancer: estimating the benefits of PSA screening. *Nat Rev Urol*. 2009;6:301–3.

31. Womble PR, Montie JE, Ye Z, Linsell SM, Lane BR, Miller DC; Michigan Urological Surgery Improvement Collaborative. Contemporary Use of Initial Active Surveillance Among Men in Michigan with Low-risk Prostate Cancer. *Eur Urol*. 2015;67:44–50.

32. Howard DH. Declines in prostate cancer incidence after changes in screening recommendations. *Arch Intern Med.* 2012;172:1267–8.

Table 1. Parameters Assigned to 3 MISCAN-based Models of PSA Screening and Treatment

	Base Model (Heijnsdijk et al[3])		Recommended Good Practice Model		Common Practice Model	
	Parameter	Source	Parameter	Source	Parameter	Source
1. Ages of men screened	55–69 years	Schröder et al[14]	55–69 years	Schröder et al[14]	55–90 years	Drazer et al[7]
2. PPV of biopsy	22.7%	Schröder et al[14]	40%	Vickers et al[18]	40%	Vickers et al[18]
3. Use of AS	AS rates depending on age, T stage and Gleason score as in ERSPC for both low-risk and non–low-risk PC; about 30% for low-risk	ERSPC data	AS rates for non–low-risk PC same as base model	ERSPC data	AS rates for non–low-risk PC same as base model	ERSPC data
			90% AS for men with low-risk tumors	Assumption	9.2% AS for men with low-risk tumors	Cooperberg et al[9]
4. Rate of side-effects	Population-based rates: 6% urinary leakage problems, 30% overall sexuality problems	Sanda et al[27]	Rates as seen in high-volume centers: 5% urinary leakage problems, 19% overall sexuality problems	Vickers et al[26]	Population-based rates: 6% urinary leakage leaking problems, 30% overall sexuality problems	Sanda et al[27]

Abbreviations: AS, active surveillance; ERSPC, European Randomized Study of Screening for Prostate Cancer; MISCAN, Microsimulation Screening Analysis; PC, prostate cancer; PPV, positive predictive value; PSA, prostate-specific antigen.

Table 2. QALYs Gained By the 3 Screening and Treatment Models At Various Health States

	Utility Estimates			Quality of Life Adjustment[a] No. of life-years		
Health State[b]	Base case	More extreme	Less extreme	Base model[c]	Recommended good practice	Common practice
Screening	0.99	0.99	1	−1.6	−1.6	−1.4
Biopsy	0.90	0.885	0.94	−1.7	−0.5	−1.6
Cancer diagnosis	0.80	0.775	0.85	−0.7	−0.7	−2.1
Radiotherapy						
At 2 months after procedure	0.73	0.72	0.82	−0.2	−0.0	−2.7
At >2–12 months	0.78	0.695	0.83	−0.9	−0.2	−11.0
Radical prostatectomy						
At 2 months after procedure	0.67	0.615	0.785	−2.0	−0.6	−3.9
At >2-12 months	0.77	0.735	0.84	−6.9	−2.1	−13.7
Active surveillance	0.97	0.91	0.985	−3.2	−9.7	−1.4
Post-recovery period[d] (1–10 years after treatment)						
Overdiagnosis[e]	0.95[d,f]	0.94	0.975	−10.8	−5.6	−24.8
No overdiagnosis	0.95[d,f]	0.94	0.975	−5.5	5.5	−19.8
Palliative therapy	0.60	0.73	0.42	14.1	14.2	18.4
Terminal illness	0.40	0.40	0.32	2.6	2.6	3.2
Total number of life-years gained	Full model	Full model	Full model	73	73	78
Total number of QALYs gained	Full model	Full model	Full model	56	74	19

Abbreviation: QALYs, quality-adjusted life-years.

[a]Numbers are over the lifetimes of 1000 men aged 0–100. Minus sign indicates number of years to be subtracted from the life-years gained in order to get the QALYs gained.

[b] For a complete list of sources of the utility values and the duration of temporary health states, see Heijnsdijk et al.[3] The more and less extreme utilities used for the sensitivity analysis are assumed to be half those previously reported, to reflect the effects of a policy on a population level, rather than the effects on the individual level.
[c]The difference in life-years for each health state has been multiplied by the utility loss to calculate the adjustment for quality of life.
[d]The following utilities translate into an aggregated utility of 0.95: urinary leakage, 0.83; bowel problems, 0.71; and sexuality problems, 0.89.
[e] Overdiagnosis implies diagnosis of prostate cancer, which in a situation without screening would not have been clinically diagnosed within the lifespan of a typical man.
[f]0.96 for the recommended good practice model.

Table 3. Predicted Effects of the 3 Screening and Treatment Models, Compared to No Screening[a]

	No Screening	Base Model	Recommended Good Practice	Common Practice
Biopsies performed	313	605	407	595[b]
Negative biopsies	201	448	250	359
Cancers diagnosed	112	157	157	236
Relative reduction in prostate cancer−specific mortality	-	37%	37%	41%
Life-years gained	-	73	73	78
QALYs gained	-	56	74	19
Relative reduction in life-years gained after adjustment for quality of life	-	23%	–1%	76%

Abbreviations: QALYs, quality-adjusted life-years.
[a]Numbers are over the lifetime of 1000 men aged 0−100.
[b]Some men undergo more than one biopsy.

Table 4. QALYs Gained/Lost By Different Aspects of Practice

	Recommended Good Practice		Common Practice	
Parameter	Aspect of practice	QALYs[a]	Aspect of practice	QALYs[a]
1. Age for screening	Limit screening in older men	Same as base model	Widespread screening of older men	34.2 (–21.8)
2. Biopsy criteria	Restrictive biopsy criteria	57.2 (+1.2)	Restrictive biopsy criteria	57.2 (+1.2)
3. AS	AS for most low-risk cancers	73.2 (+17.2)	Little use of AS	49.1 (–6.9)
4. Regionalization	Most treatment at high-volume centers	59.3 (+3.3)	Much treatment at low-volume centers	Same as base model
Total	All four of the above factors	74.0 (+18.0)	All four of the above factors	19.0 (–37.0)

Abbreviations: AS, active surveillance; QALYs, quality-adjusted life-years.

[a]Number in parentheses indicates incremental/decremental effect on QALYs as compared to base model's 56.0 QALYs gained.

Supplementary material. Sensitivity Analysis: Effects of Various Modeling Assumptions on Total QALYs Gained

	Base model		Recommended Good Practice		Common Practice	
	Scenario 5[e]		Scenario 5[e]		Scenario 5[e]	
	Terminal illness 0.4	Terminal illness 0.32	Terminal illness 0.4	Terminal illness 0.32	Terminal illness 0.4	Terminal illness 0.32
Scenario 1[a]	39.8	40.1	49.7	50.0	–7.8	–7.4
Scenario 2[b]	77.6	77.9	88.5	88.8	61.1	61.5
Scenario 3[c]	60.5	60.9	55.7	56.1	42.9	43.4
Scenario 4[d]	55.2	55.6	80.7	81.1	5.2	5.6

In the base model, screening men between the ages of 55–69 years yields 56 QALYs gained over a lifetime. This is based on assumptions of the utilities for the modeled health states; screening attendance, biopsy, diagnosis, treatment, post-recovery period, palliative treatment, and terminal illness. These utilities can be varied from less extreme to more extreme values (Table 2).

[a]Treatment and procedures less tolerable (ie, low utilities for everything but terminal illness)
[b]Treatment and procedures more tolerable (ie, high utilities for everything but terminal illness)
[c]Cancer worry less tolerable, treatment side effects more tolerable (ie, low utilities for active surveillance, biopsy, diagnosis, and screening; high utilities for treatment and recovery)
[d]Vice versa of Scenario 3 (ie, high utilities for active surveillance, biopsy, diagnosis and screening; low utilities for treatment and recovery)
[e]Scenarios 1–4, with different utility value for terminal illness