The ONCOTYROL Prostate Cancer Outcome and Policy Model: Effect of Prevalence Assumptions on the Benefit-Harm Balance of Screening

Nikolai Mühlberger, DVM, MPH, Christina Kurzthaler, MSc, Rowan Iskandar, MA, Murray D. Krahn, MD, MSc, Karen E. Bremner, BSc, Willi Oberaigner, Priv.-Doz., Dr., Helmut Klocker, Mag., PhD, Wolfgang Horninger, MD, Annette Conrads-Frank, PhD, Gaby Sroczynski, MPH, Dr.PH, Uwe Siebert, MD, MPH, MSc, ScD

Background. *The ONCOTYROL Prostate Cancer Outcome and Policy (PCOP) model is a state-transition microsimulation model evaluating the benefits and harms of prostate cancer (PCa) screening. The natural history and detection component of the original model was based on the 2003 version of the Erasmus MIcrosimulation SCreening ANalysis (MISCAN) model, which was not calibrated to prevalence data. Compared with data from autopsy studies, prevalence of latent PCa assumed by the original model is low, which may bias the model toward screening. Our objective was to recalibrate the original model to match prevalence data from autopsy studies as well and compare benefit-harm predictions of the 2 model versions differing in prevalence.* ***Methods***. *For recalibration, we reprogrammed the natural history and detection component of the PCOP model as a deterministic Markov state-transition cohort model in the statistical software package R. All parameters were implemented as variables or time-dependent functions and calibrated simultaneously in a single run. Observed data used as calibration targets included data from autopsy studies, cancer registries, and the European Randomized Study of Screening for Prostate Cancer. Compared models were identical except for calibrated parameters.* ***Results***. *We calibrated 46 parameters. Prevalence from autopsy studies could not be fitted using the original parameter set. Additional parameters, allowing for interruption of disease progression and age-dependent screening sensitivities, were needed. Recalibration to higher prevalence demonstrated a considerable increase of overdiagnosis and decline of screening sensitivity, which significantly worsened the benefit-harm balance of screening.* ***Conclusions***. *Our calibration suggests that not all cancers are at risk of progression, and screening sensitivity may be lower at older ages. PCa screening models that use calibration to simulate disease progression in the unobservable latent phase are highly sensitive to prevalence assumptions.* ***Key words:*** *decision analysis; Markov models; simulation methods; prostate cancer; cancer prevention.* ***(Med Decis Making 2015;35:758–772)***

Prostate cancer (PCa) is the most frequently diagnosed male malignancy and the third most frequent cause of male cancer death in the World Health Organization (WHO) European Region, including Austria.[1] Annually, about 5000 of the approximately 4.1 million Austrian men are newly diagnosed with PCa, and around 1100 die of PCa.[2]

PCa mortality may be reduced by early detection and treatment. In the Austrian state of Tyrol, a 30% reduction in PCa mortality has been observed since the introduction of prostate-specific antigen (PSA) screening for men aged 45 to 74 years in 1993.[3–5] The European Randomized Study of Screening for Prostate Cancer (ERSPC) showed reductions in PCa mortality (21%–29%)[6] and metastatic disease (30%)[7] after median follow-ups of 11 and 12 years, respectively. In contrast, other trials, including the Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer screening trial[8] and a meta-analysis of trials,[9]

DOI: 10.1177/0272989X15585114

did not show significant reductions in PCa-specific mortality.

Despite inconclusive evidence on mortality reduction, the main concern with PCa screening is that gains in lifetime and quality of life (QoL) achieved by prevention of deaths and metastatic disease might be offset by opposing losses in QoL due to overdiagnosis and overtreatment. QoL may be impaired by diagnostic biopsies; treatment with radical prostatectomy (RP), radiotherapy (RT), and androgen deprivation therapy (ADT); and the frequent long-term adverse events of treatment, including erectile dysfunction (ED), urinary incontinence (UI), and bowel dysfunction (BD). In addition, RP has a small risk of perioperative death.[10,11]

QoL losses are particularly harmful for overdiagnosed men, whose cancer would remain undetected for their remaining lifetime without screening. As a substantial fraction of PCas shows late onset and slow progression, overdiagnosis is a serious problem of PCa screening, which increases when screening is performed repeatedly or in men with relatively short life expectancy. The magnitude of overdiagnosis is difficult to observe. However, modeled estimates from the Erasmus MIcrosimulation SCreening ANalysis (MISCAN) model suggest that the percentage of screen-detected cancers that would remain undetected without screening could be up to 50%.[12–14]

Due to uncertainty about its benefit-harm balance, some medical societies in Europe and the United States, including the European Association of Urology and the United States Preventive Services Task Force (USPSTF), do not recommend routine PSA-based screening.[15,16]

More than 25 mathematical models have been developed since the early 1990s to simulate the effects of PSA screening.[14,17–40] However, only 8 models considered QoL,[14,18,22,23,31,36,37,40] of which only 3 accounted for overdiagnosis.[14,23,31] Six of the 8 models indicated that the benefits of screening can be outweighed by its harms in general or depending on screening frequency and age.[18,22,23,31,37,40] However, a recent version of the Erasmus MISCAN model predicted gains in quality-adjusted life years (QALYs) even with annual screening up to age 74 years and clearly supports screening.[14]

The Erasmus MISCAN model is based on data of the ERSPC and can be considered the most comprehensive model for the European context. A first version of the model simulating the development of PCa up to the time of detection was published by Draisma and others[12] in 2003, which since then has been stepwise extended and updated.[14,41,42] The model is distinctive in that it explicitly models the natural history of PCa, including its onset and progression in stage and grade in the unobservable latent phase, which was achieved by calibration to observed data from cancer registries and the ERSPC.

Screening-related gains in lifetime and QoL were also predicted by the original version of the Prostate Cancer Outcome and Policy (PCOP) model, which was developed within the Tyrolean ONCOTYROL translational research cooperation based on structural elements and calibrated natural history and detection parameters of the 2003 version of the MISCAN model. However, benefit-harm predictions of the original PCOP model may be biased, since parameters adopted from the MISCAN model were not calibrated to data on prevalence of latent cancer. Matching age-specific latent prevalence, however, seems important for screening models that are calibrated to fit observed clinical incidence, as it may affect the risk of overdiagnosis and the sensitivity of screening, which both can influence the predicted benefit-harm balance of screening. In addition, the latent prevalence pool is the target for all screening measures and therefore should be estimated correctly. Comparing latent prevalence predicted by the original PCOP model with data from

Received 8 August 2014 from the Department of Public Health and Health Technology Assessment, UMIT–University for Health Sciences, Medical Informatics and Technology, Tyrol, Austria (NM, CK, RI, ACF, GS, US); Division of Health Technology Assessment and Bioinformatics, ONCOTYROL–Center for Personalized Cancer Medicine, Innsbruck, Austria (NM, CK, RI, ACF, GS, US); Department of Health Policy and Management, University of Minnesota, Minneapolis, MN, USA (RI); Toronto Health Economics and Technology Assessment (THETA) Collaborative, University of Toronto, Toronto, Ontario, Canada (MDK); Toronto General Research Institute, Toronto General Hospital, Toronto, Ontario, Canada (MDK, KEB); Cancer Registry of Tyrol, TILAK GmbH, Innsbruck, Austria (WO); Department of Urology, Innsbruck Medical University, Innsbruck, Austria (HK, WH); Center for Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA (US); and Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA (US). Financial support for this study was provided entirely by the COMET Center ONCOTYROL, which is funded by the Austrian Federal Ministries BMVIT/BMWFJ (via FFG) and the Tiroler Zukunftsstiftung/Standortagentur Tirol (SAT). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Revision accepted for publication 6 April 2015.

Address correspondence to Uwe Siebert, MD, MPH, MSc, ScD, UMIT–University for Health Sciences, Medical Informatics and Technology, Eduard Wallnoefer Center I, A-6060 Hall i.T., Tirol, A-6060, Austria; telephone: +43-50-8648-3930; fax: +43-50-8648-673930; e-mail: public-health@umit.at.

autopsy studies[43–48] suggests a considerable underestimation of the latent prevalence pool, given that cancers detected at autopsy are biopsy detectable as well.

Our objective was to investigate how predicted benefits and harms of PCa screening are affected by the size of the latent prevalence pool. Under the assumption that cancers detected at autopsy would also be detectable by biopsy, we recalibrated all natural history and detection parameters of the PCOP model adopted from the 2003 MISCAN model to match prevalence data from autopsy studies as well. Benefits and harms of screening predicted by the recalibrated PCOP model were then contrasted to predictions of the original model version.

The primary purposes of this work are to present the ONCOTYROL PCOP model, to report findings from the complex calibration of natural history and detection parameters, and to show the effect of the size of the prevalence pool on the benefit-harm balance of screening. We do not intend to present a comprehensive benefit-harm evaluation of screening, which will be the topic of a subsequent publication.

METHODS

Model Structure

The ONCOTYROL PCOP model is a decision-analytic state-transition microsimulation model programmed using the software TreeAge Pro 2012 (TreeAge Software Inc., Williamstown, MA). It simulates the natural history of PCa from onset to death and the lifetime consequences of screening and treatment on duration and quality of life following the International Society For Pharmacoeconomics and Outcomes Research–Society for Medical Decision Making (ISPOR-SMDM) guidelines on state transition modeling.[49]

The structure of the ONCOTYROL PCOP model is depicted in Figure 1. Men are born without cancer, and during their lifetime, they may or may not develop latent cancer, which can progress in stage and grade. Distinguished tumor stages are localized cancer (corresponding to Union for International Cancer Control TNM classification T1/2 N0/X M0/X), regional cancer (TNM classification T3+ or N+ with M0/X), and distant cancer (TNM classification any TN and M1). Each cancer stage is separated into 3 states according to Gleason score grading ($G < 7$, $G = 7$, and $G > 7$), which yields a total of 9 stage- and grade-specific cancer states. Latent cancer can be detected by clinical symptoms or at an earlier stage through screening (e.g., PSA). Once detected, cancer can be treated. Treatment choice (e.g., RP, RT, or ADT) and effectiveness depend on cancer stage and may cause specific long-term complications (e.g., ED, UI, and BD). Treatment of early stages can result in cure. Without cure, cancer may progress and eventually kill the patient, unless he dies from another cause before.

The natural history and detection component of the model is designed as a state-transition model that can be analyzed with deterministic cohort simulation and microsimulation. Health states and transitions for this initial part of the model are adopted from the 2003 version of the Erasmus MISCAN model by Draisma and others.[12] Cancer states are modeled as tunnel states to keep track of the time already spent in the current health state. Events following cancer detection are modeled via microsimulation only. In this part of the model, individual patient history is recorded and directed via tracker variables to limit the number of health states. There were 2 important reasons for integrating parts of the Erasmus MISCAN model into our model. First, it is fitted to data of the ERSPC, which currently can be considered to provide the best available evidence on the effectiveness of PCa screening. Second, unlike most other models, the Erasmus model simulates the complete natural history of cancer from latent onset to death in the absence and presence of screening. Thus, it provides a straightforward way to estimate the lifetime risk of overdiagnosis, which, assuming that only clinically detected cancers are relevant, is the difference in lifetime risk of cancer diagnosis with and without screening.

Natural History and Detection Data (Calibration)

Since cancer onset and progression in the latent phase are unobservable, parameters for this part of the model have to be calibrated. The original version of the PCOP model applied parameters from the 2003 Erasmus MISCAN model calibrated to data from cancer registries and the ERSPC.[12] We recalibrated the parameters to also match prevalence data from autopsy studies, following the ISPOR-SMDM guidelines on parameter estimation.[50]

For recalibration, we reprogrammed the natural history and detection component of the PCOP model using the statistical software package R (www.r-project.org) as a deterministic Markov state-transition cohort model with annual cycle length and parameters implemented as functions or variables. Functions were used for exit probabilities from the noncancer and cancer states, which determine what proportion of men exits the states during a time cycle.

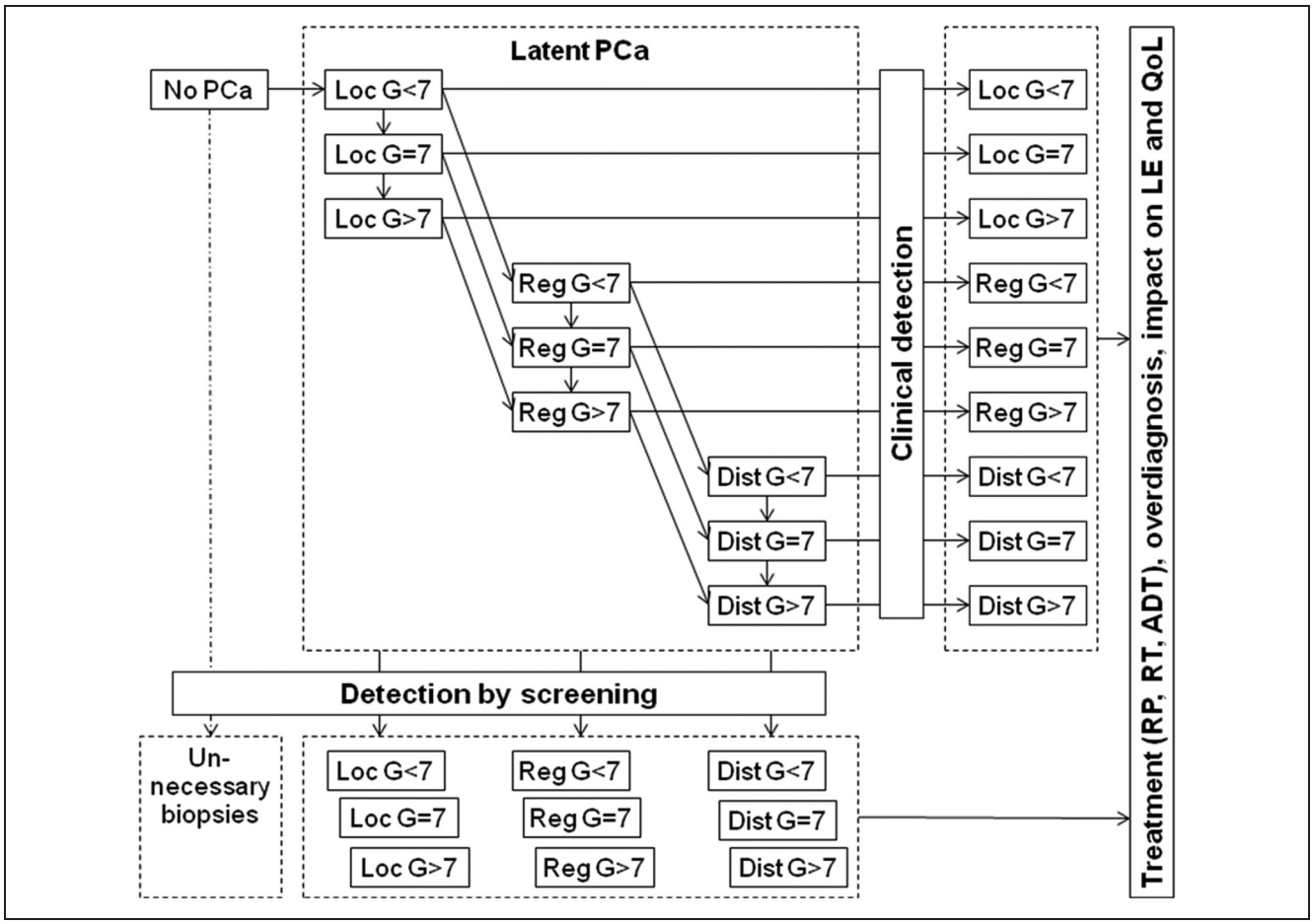

Figure 1 Structure of the ONCOTYROL Prostate Cancer Outcome and Policy model. The model simulates the natural history of prostate cancer from onset to death and the consequences of screening and treatment on duration and quality of life. Cancer mortality is considered in clinical distant cancer states. Death due to other causes can occur in all health states. The natural history and detection component of the model is designed as a state-transition model that can be analyzed with deterministic cohort simulation and microsimulation. Events following cancer detection summarized at the right end of the diagram are modeled via microsimulation only. In that part of the model, individual patient history is recorded and directed via tracker variables. PCa, prostate cancer; Loc, localized cancer (T1/2, N0/X, M0/X); Reg, regional cancer (T3/4 or N+ and M0/X); Dist, distant cancer (any TN, M1); G, Gleason score; RP, radical prostatectomy; RT, radiotherapy; ADT, androgen deprivation therapy; LE, life expectancy; QoL, quality of life.

To which health state men transit is determined by consecutive probabilities implemented as variables. For exit probabilities, we assumed 1) that the risk of developing cancer or progressing in tumor stage or grade may depend on time spent in the current state and 2) that not all men would progress to cancer even if they lived forever.

On the basis of these assumptions, we implemented a Weibull distributed density function weighted by the proportion of men at risk to derive time-dependent hazard rates (see equation (1)). An extension of the equation was used to convert the rates into annual probabilities required by our state-transition model.

$$h(t) = \frac{p\left(\frac{a}{b}\left(\frac{t}{b}\right)^{a-1} e^{-\left(\frac{t}{b}\right)^{a}}\right)}{1 - p\left(1 - e^{-\left(\frac{t}{b}\right)^{a}}\right)}, \quad (1)$$

where

p = proportion of men at risk of exiting state (i.e., to latent cancer or next cancer state),
a = Weibull shape parameter,
b = Weibull scale parameter, and
t = time already spent in current health state.

As in the MISCAN model, screening was considered as a single test consisting of PSA testing

followed by ultrasound-guided sextant biopsy, when PSA level is $\geq$3 ng/mL. Sensitivities of the test combination were calibrated using parameters allowing sensitivities to vary with tumor stage and age.

All parameters were calibrated simultaneously in a single run using the ''nlminb'' optimization algorithm available in R to minimize the relative difference between observed and predicted data ($|observed - predicted|/observed$). As calibration targets, we used age-specific latent prevalence data from autopsy studies from the prescreening era, which were used by Etzioni and others[45] to calculate the incidence of latent PCa, and data applied for the calibration of the 2003 MISCAN model.[12] Those included age-specific cancer incidence and stage distribution of clinically detected cancers from the prescreening era supplied by Dutch cancer registries, as well as age-specific detection rates and stage and grade distribution of cancers detected in the first screening round of the Rotterdam center of the ERSPC trial. To account for background mortality in the prescreening and trial period, we applied mortality data for Dutch men from 1991 and 2000 from Eurostat.[51]

Not included in our calibration, but taken from the literature, were parameters for PCa-specific mortality and screening characteristics needed to adjust the number of biopsies to include those with false-negative results and those performed unnecessarily due to false-positive PSA tests. Specificity of PSA testing and sensitivity and specificity of biopsy required for the adjustment were assumed to be 85%,[52] 90%,[53,54] and 100%, respectively. Screening adherence was assumed to be 100%. PCa mortality was modeled based on 3-year conditional relative survival data for distant PCa reported by the Surveillance, Epidemiology, and End Results (SEER) program.[55] No increase in mortality was assumed from local and regional cancer.

All parameters applied in the natural history and detection component of the recalibrated PCOP model are summarized in the first 2 sections of Table 1. Calibrated parameters applied in the original PCOP are reported by Draisma and others.[12]

Data on Effectiveness and Adverse Events of Treatment

Beneficial and harmful outcomes related to treatment were modeled based on published data reported in the third section of Table 1. The PCOP model distinguishes treatment of local, regional, and distant cancer, each of which may consist of various treatment options. In this work, we do not consider treatment heterogeneity but assume that all detected cancers receive homogeneous stage-specific treatment consisting of RP, RT, and ADT for localized, regional, and distant cancer, respectively, which largely reflects treatment patterns observed by the Tyrolean early PCa detection program from 1993 to 2005.[56] Benefits of screening mainly result from cancer cure achievable with early treatment. Our model applies grade-specific cure rates for local and regional treatment calibrated by Wever and others[57] to the 27% mortality reduction observed in ERSPC at the 9-year follow-up.[58] Treatment for distant cancer was assumed to have no effect on mortality.

Adverse events of local or regional cancer treatment considered by our model are ED, UI, BD, and perioperative mortality of RP. Since adverse events due to RP and RT differ in frequency, severity, and duration, these were modeled separately. Frequency of adverse events was modeled using attributable risks from a meta-analysis of studies comparing event risks in treated patients and patients under active surveillance.[11] Severity and duration of adverse events were derived from posttreatment PCI (UCLA prostate cancer index) functional scores assessed in the ERSPC.[14,59] Adverse events of ADT were not modeled explicitly but covered by the utility weight assigned to the distant cancer states.

Utility Data

The utilities applied in the PCOP model are presented in the last section of Table 1. Based on the assumption that nonmetastatic PCa does not significantly affect quality of life, we used age-specific utilities reported for the general male population in Sweden,[60] which we consider to closely reflect Austrian utilities, to weight lifetime in nonmetastatic and preclinical metastatic cancer states. For time in clinical metastatic cancer states, we applied a weight of 0.6, which accounts for disease- and treatment-related disutility.[14]

QoL losses due to treatment-related sexual, urinary, and bowel dysfunction were taken into account by multiplying the utilities for nonmetastatic cancer states by PCI score-specific utilities derived from a Canadian study.[61] These multiplicative disutilities were assigned only to patients who experienced a treatment-related dysfunction for an assumed 5-year duration.

Additional QoL impairments assumed to last for 1 year at most were implemented as one-time QALY decrements subtracted from a man's quality-adjusted lifetime. Decrements were used for biopsy, RP, RT, and the terminal cancer phase and calculated based

Table 1 Parameters of the Recalibrated ONCOTYROL PCOP Model

Parameters	Base-Case Values	Source
Natural history		
Probability of exiting the no cancer state (p/scale/shape)[a]	0.838/80.427/8.448	Calibrated[b]
Probability of exiting local G < 7 cancer state (p/scale/shape)[a]	0.449/2.041/8.431	Calibrated[b]
Probability of exiting local G = 7 cancer state (p/scale/shape)[a]	0.811/1.292/4.349	Calibrated[b]
Probability of exiting local G > 7 cancer state (p/scale/shape)[a]	0.987/2.940/7.069	Calibrated[b]
Probability of exiting regional G < 7 cancer state (p/scale/shape)[a]	0.450/6.050/4.129	Calibrated[b]
Probability of exiting regional G = 7 cancer state (p/scale/shape)[a]	0.560/4.113/5.546	Calibrated[b]
Probability of exiting regional G > 7 cancer state (p/scale/shape)[a]	0.823/2.024/2.791	Calibrated[b]
Probability of exiting distant G < 7 cancer state (p/scale/shape)[a]	0.999/0.254/5.373	Calibrated[b]
Probability of exiting distant G = 7 cancer state (p/scale/shape)[a]	0.945/0.806/4.564	Calibrated[b]
Probability of exiting distant G > 7 cancer state (p/scale/shape)[a]	0.999/1.135/5.521	Calibrated[b]
Probability of local G < 7 cancer progress to regional	0.158	Calibrated[b]
Probability of local G = 7 cancer progress to regional	0.388	Calibrated[b]
Probability of regional G < 7 cancer progress to distant	0.005	Calibrated[b]
Probability of regional G = 7 cancer progress to distant	0.144	Calibrated[b]
Probability of dying of PCa conditional on survival	SEER data	National Cancer Institute[55]
Age-specific probability of dying of other causes	Austrian life table 2010/12	Statistik Austria[62]
Cancer detection (clinically or by screening)		
Probability of local G < 7 cancer to be clinically detected	0.006	Calibrated[b]
Probability of local G = 7 cancer to be clinically detected	0.110	Calibrated[b]
Probability of local G > 7 cancer to be clinically detected	0.604	Calibrated[b]
Probability of regional G < 7 cancer to be clinically detected	0.067	Calibrated[b]
Probability of regional G = 7 cancer to be clinically detected	0.108	Calibrated[b]
Probability of regional G > 7 cancer to be clinically detected	0.407	Calibrated[b]
Probability of distant G < 7 cancer to be clinically detected	0.233	Calibrated[b]
Probability of distant G = 7 cancer to be clinically detected	0.897	Calibrated[b]
Probability of distant G > 7 cancer to be clinically detected	1.000	Assumption
Probability of participating in screening	1	Assumption
Probability of detecting local cancer by screening (age <70 y)	0.550	Calibrated[b]
Probability of detecting local cancer by screening (age 70+ y)	0.370	Calibrated[b]
Probability of detecting regional/distant PCa by screening (age <70 y)	0.677	Calibrated[b]
Probability of detecting regional/distant PCa by screening (age 70+ y)	0.456	Calibrated[b]
Specificity of PSA (to account for disutility by unnecessary biopsies)	0.85	Wolf et al.[52]
Sensitivity of biopsy (to account for disutility by false-negative biopsies)	0.90	Fink et al.[53]; Shariat and Roehrborn[54]
Specificity of biopsy	1	Assumption

(continued)

Table 1 (continued)

Parameters	Base-Case Values	Source
Treatment (beneficial and harmful events)		
Probability of cure given local/regional cancer (G < 7)	0.42	Wever et al.[57]
Probability of cure given local/regional cancer (G > 7)	0.23	Wever et al.[57]
Probability of cure given distant cancer (all G)	0	Wever et al.[57]
Risk of dying from prostatectomy (30-day mortality)	0.0015	Wolf et al.[52]
Risk of erectile dysfunction attributable to prostatectomy	0.28	Chou et al.[11]
Risk of erectile dysfunction attributable to radiotherapy	0.15	Chou et al.[11]
Risk of urinary incontinence attributable to prostatectomy	0.22	Chou et al.[11]
Risk of urinary incontinence attributable to radiotherapy	0.031	Chou et al.[11]
Risk of bowel dysfunction attributable to prostatectomy	0	Chou et al.[11]
Risk of bowel dysfunction attributable to radiotherapy	0.028	Chou et al.[11]
Duration of treatment related dysfunctions	5 years	Korfage et al.[59]
Utilities		
Utility without clinical distant PCa and treatment complication	Age specific	Olsen et al.[60]
Utility of clinical distant cancer	0.6	Heijnsdijk et al.[14]
Utility of erectile dysfunction by RP (PCI score 0–25)	0.89	Heijnsdijk et al.[14]; Korfage et al.[59]; Krahn et al.[61]
Utility of erectile dysfunction by RT (PCI score >25–50)	0.95	Heijnsdijk et al.[14]; Korfage et al.[59]; Krahn et al.[61]
Utility of urinary incontinence by RP (PCI score >50-75)	0.90	Heijnsdijk et al.[14]; Korfage et al.[59]; Krahn et al.[61]
Utility of urinary incontinence by RT (PCI score >75–100)	0.93	Heijnsdijk et al.[14]; Korfage et al.[59]; Krahn et al.[61]
Utility of bowel dysfunction by RP (PCI score >75–100)	0.93	Heijnsdijk et al.[14]; Korfage et al.[59]; Krahn et al.[61]
Utility of bowel dysfunction by RT (PCI score >75–100)	0.93	Heijnsdijk et al.[14]; Korfage et al.[59]; Krahn et al.[61]
One-time utility decrement due to biopsy	–0.006 QALY	Calculated from Heijnsdijk et al.[14]
One-time utility decrement due to RP	–0.214 QALY	Calculated from Heijnsdijk et al.[14]
One-time utility decrement due to RT	–0.196 QALY	Calculated from Heijnsdijk et al.[14]
One-time incremental utility decrement due to terminal PCa	–0.099 QALY	Calculated from Heijnsdijk et al.[14]

PCa, prostate cancer; PCI, prostate cancer index; PCOP, Prostate Cancer Outcome and Policy; PSA, prostate-specific antigen; QALY, quality-adjusted life year; RP, radical prostatectomy; RT, radiotherapy; SEER, Surveillance, Epidemiology, and End Results.

a. Parameters for equation (1).

b. Calibrated parameters applied in the original PCOP are published by Draisma and others.[12] (Note that the original PCOP model does not use parameter p for exit probabilities, except for exit from the no cancer state.)

Table 2 Calibration—Observed and Predicted Stage and Grade Distribution

	Observed (%)				Predicted by Recalibrated PCOP Model (%)				Predicted by Original PCOP Model (%)			
	G < 7	G = 7	G > 7	All	G < 7	G = 7	G > 7	All	G < 7	G = 7	G > 7	All
Comparison with cancer registry data from prescreening era												
Local PCa				58.03	1.67	17.82	38.55	58.03	21.66	21.75	15.44	58.85
Regional PCa				18.81	0.98	3.50	14.33	18.81	2.00	4.21	12.79	18.99
Distant PCa				23.15	0.00	3.85	19.30	23.15	0.08	5.25	16.84	22.16
All				100.00	2.65	25.17	72.18	100.00	23.73	31.20	45.07	100.00
Comparison with ERSPC screening arm (first screening round)												
Local PCa	56.96	17.41	4.45	78.82	56.96	17.41	4.45	78.82	56.93	17.46	4.17	78.56
Regional PCa	7.93	9.77	2.90	20.60	7.93	9.77	2.90	20.60	8.08	9.80	3.05	20.92
Distant PCa	0.00	0.10	0.48	0.58	0.01	0.09	0.48	0.58	0.00	0.09	0.43	0.52
All	64.89	27.28	7.83	100.00	64.90	27.27	7.83	100.00	65.01	27.35	7.64	100.00

ERSPC, European Randomized Study of Screening for Prostate Cancer; G, Gleason score; PCOP, Prostate Cancer Outcome and Policy.

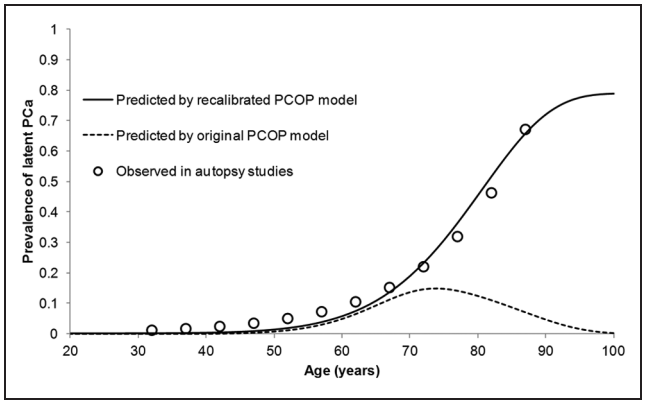

Figure 2 Calibration—observed and predicted age-specific prevalence of latent prostate cancer (PCa). PCOP, Prostate Cancer Outcome and Policy. (Source of observed prevalence: Etzioni and others.[45])

on data and assumptions of the 2012 version of the Erasmus MISCAN model.[14]

Comparative Model Analyses

To assess the impact of considering the latent prevalence pool on the benefits and harms of PCa screening, we compared outputs of the recalibrated PCOP model with outputs of the original model using parameters of the 2003 MISCAN model. Both models were identical except for calibrated parameters and cycle length, which is yearly for the recalibrated and monthly for the original model. The reason for shorter time cycles of the original model is that the MISCAN model is a discrete-event model. Implementing parameters from that type of model in a state-transition model may require modeling short time cycles to minimize cycle length–related competing risk effects and thus to achieve concordant model outputs.

Strategies evaluated for the model comparison were no screening; one-time screening at ages 55, 65, and 75 years; and screening at intervals of 4 years, 2 years, and 1 year at ages 55 to 74 years. The strategies were chosen deliberately to show the effects of age and screening frequency. Adherence with screening and treatment was assumed to be 100% (i.e., actual screening rather than intention to screen) to eliminate the effect of external behavioral factors. Models were analyzed by microsimulation modeling the history of 6 million men from birth to death. Background mortality was modeled using life table data for Austrian men from 2010/2012.[62] Model outputs were chosen to cover the broad spectrum of screening-related benefits and harms.

RESULTS

Calibration of Natural History and Detection Parameters

We calibrated 46 parameters. Parameter values are given in Table 1. Agreement of model predictions with observed data is presented in Figures 2 to 4 and Table 2. A comparison of the recalibrated and original models indicates that clinical incidence from the prescreening era is better fitted by the original model, whereas stage and grade distributions and detection rates are better fitted by the recalibrated model.

Prevalence of latent cancer was not considered a calibration target for parameters of the original model. Therefore, the recalibrated model provides a better match of age-specific latent prevalence data from autopsy studies. Areas under the curves in Figure 2 indicate that the latent prevalence pool assumed by the recalibrated model is about 4-fold larger. Since prevalence of latent PCa is mainly

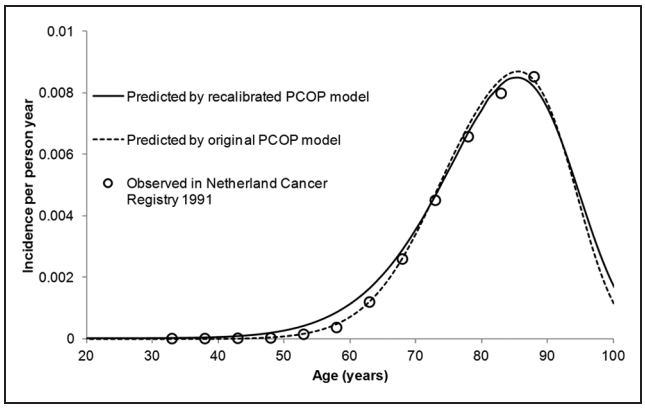

Figure 3 Calibration—observed and predicted incidence of prostate cancer in the prescreening era. PCOP, Prostate Cancer Outcome and Policy.

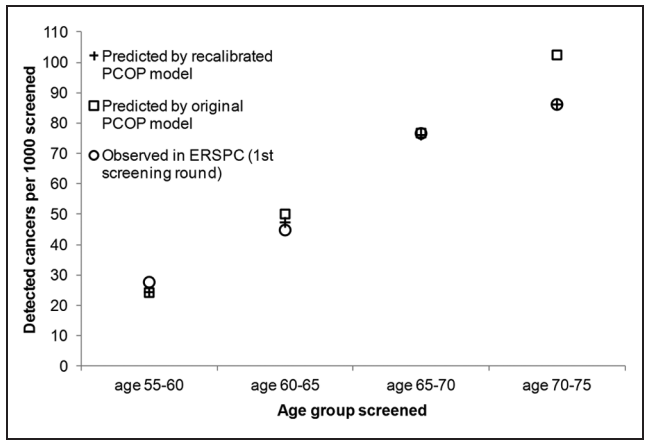

Figure 4 Calibration—observed and predicted detection rates in the European Randomized Study of Screening for Prostate Cancer (ERSPC) Rotterdam. PCOP, Prostate Cancer Outcome and Policy.

determined by entries into the prevalence pool, the predicted lifetime risk of developing latent cancer increased from 15% to 38%.

An important finding of the calibration was that prevalence data from autopsy studies could not be fitted by the parameters used in the calibration of the 2003 MISCAN model. Additional parameters had to be introduced to compensate for the increase in latent prevalence.

In particular, parameters allowing for an interruption of disease progression in the stage- and grade-specific cancer states had to be introduced to reach an acceptable match of observed clinical incidence data. In our model, these parameters are represented by the proportion of men at risk of exiting a cancer state used as weights for the time-dependent hazard function (see parameter p in equation (1)). Values calibrated for the at-risk proportion increase with stage and grade. For example, in the local $G < 7$ state, it is 45%, whereas in the local $G > 7$ or distant states, it approaches 100%.

Fitting of age-specific detection rates observed in the ERSPC trial required implementation of an effect modifier allowing for lower screening sensitivities above age 70 years. Applying stage-specific sensitivities of screening calibrated by Draisma and others,[12] which were 64%, 91%, and 97% for local, regional, and distant cancer, respectively, yielded a good match of detection rates in men younger than 70 years but not in older men, in whom the detection rate was overestimated about twice. Screening sensitivities for local and higher cancer states calibrated by our model with higher prevalence were 55% and 68% for screening below age 70 years, respectively, and 37% and 46% thereafter.

Model Results and Comparison

Outputs of the recalibrated and original versions of the PCOP model are contrasted in upper and lower sections of Table 3, respectively.

Outputs of the recalibrated model suggest that the lifetime risk of cancer diagnosis increases with age and screening frequency, from 9% without screening to 25% with annual interval screening. However, lifetime risks for screening-related diagnosis and overdiagnosis increase in parallel up to 21.4% and 16.5%, respectively. Percentages of overdiagnosis in screen-detected cancers range from 77% with annual screening to 87% with one-time screening at age 75.

All screening strategies are predicted to reduce PCa-specific mortality. Without screening, the lifetime risk of PCa death is estimated to be 1.64%. Mortality reductions by screening range from 2% with one-time screening at age 55 years to 36% with annual screening. Mortality reduction by one-time screening was highest at age 65 years (6%). Lifetime gains per man range from 1 day with one-time screening at age 75 years to 21 days with annual screening. One-time screening is most effective at age 65 years, gaining 3 days.

The benefit-harm balance of all screening strategies becomes negative when utility is considered. Expected losses in quality-adjusted life days (QALDs) per man range from 1 QALD with early one-time screening to 14 QALDs with annual screening, increasing with age and screening frequency. In our model, the benefits of screening and early treatment are offset by perioperative mortality of RP and QoL losses due to biopsy, treatment, and treatment-

Table 3 Model Comparison: Effect of Recalibration on Benefits and Harms of PCa Screening

	No Screening	One-Time Screening at 55 y	One-Time Screening at 65 y	One-Time Screening at 75 y	Interval Screening Every 4 y, 55–74y (5 Times)	Interval Screening Every 2 y, 55–74 y (10 Times)	Interval Screening Every 1 y, 55–74 y (20 Times)
Results of recalibrated PCOP model (large prevalence pool)							
Lifetime risk of PCa diagnosis (%)	8.99	10.17	13.18	15.39	18.81	22.67	25.45
Lifetime risk of PCa diagnosis by screening (%)	—	1.47	5.05	7.39	12.50	17.70	21.42
Lifetime risk of overdiagnosis (%)	—	1.18	4.19	6.40	9.82	13.68	16.46
Overdiagnosis in screen-detected PCa (%)	—	80.27	82.97	86.60	78.56	77.29	76.84
Lifetime risk of dying of PCa (%)	1.64	1.61	1.55	1.57	1.35	1.18	1.06
Lifetime gained v. no screening (days)	—	1.25	3.07	1.16	10.60	15.81	21.02
QALDs gained v. no screening	—	−0.98	−3.72	−7.63	−6.96	−10.21	−13.61
RP-related deaths per 10,000 men (*n*)	0.71	0.90	1.23	1.42	2.03	2.61	3.36
RP- and RT-related AEs per man (*n*)	0.03	0.03	0.05	0.06	0.08	0.10	0.12
PSA tests per man (*n*)	—	0.93	0.82	0.62	4.02	7.68	14.86
False-positive PSA tests per man (*n*)	—	0.13	0.11	0.06	0.57	1.10	2.16
PSA tests needed to avoid 1 death (*n*)	—	2562	882	838	1343	1648	2539
Results of original PCOP model (small prevalence pool)							
Lifetime risk of PCa diagnosis (%)	8.32	8.57	10.12	11.60	13.07	14.46	15.14
Lifetime risk of PCa diagnosis by screening (%)	—	1.35	5.30	6.91	11.60	13.90	14.86
Lifetime risk of overdiagnosis (%)	—	0.25	1.8	3.28	4.75	6.14	6.82
Overdiagnosis in screen-detected PCa (%)	—	18.52	33.96	47.47	40.95	44.17	45.90
Lifetime risk of dying of PCa (%)	2.55	2.30	2.01	2.19	1.52	1.41	1.40
Lifetime gained v. no screening (days)	—	10.21	13.30	7.28	26.37	26.98	27.58
QALDs gained v. no screening	—	7.69	6.68	0.01	9.93	6.19	2.61
RP-related deaths per 10,000 men (*n*)	0.69	0.82	1.10	1.09	1.91	2.10	2.20
RP- and RT-related AEs per man (*n*)	0.03	0.04	0.05	0.05	0.07	0.08	0.09
PSA tests per man (*n*)	—	0.93	0.84	0.65	4.10	7.88	15.39
False-positive PSA tests per man (*n*)	—	0.14	0.11	0.08	0.59	1.15	2.27
PSA tests needed to avoid 1 death (*n*)	—	374	154	177	395	693	1332

Results are based on individual-level simulation (microsimulation) with 6 million trials. Time horizon = 120 years, compliance = 100%. Recalibrated model with annual cycles, original model with monthly cycles. AE, adverse event; PCa, prostate cancer; PCI, prostate cancer index; PCOP, Prostate Cancer Outcome and Policy; PSA, prostate-specific antigen; QALD, quality-adjusted life day; RP, radical prostatectomy; RT, radiotherapy.

related adverse events. RP-related mortality is predicted to range from 0.7 deaths per 10,000 men without screening to 3.4/10,000 men with annual screening. The expected number of other adverse events per man due to RP and RT ranges from 0.03 without screening to 0.12 with annual screening.

The number of PSA tests decreases as the screening interval and age at screening increase. Estimates for one-time screening suggest that moving from one-time screening at age 55 years to one-time screening at age 75 years reduces the number of tests by 33%. Moving from annual to quadrennial screening reduces the number of PSA tests by 73%. Frequency of false-positive PSA results, which trigger unnecessary biopsies, and number of PSA tests needed to prevent 1 PCa death are similarly related to age and screening interval. With annual screening, a man would have false-positive results twice in his life, whereas with one-time screening at age 75 years, the risk is only 6%. The number of PSA tests needed to prevent 1 PCa death was highest with one-time screening at age 55 years (2562) and annual interval screening (2539).

In contrast to the recalibrated model's prediction of QALY losses with screening, the original model is clearly in favor of screening. It predicts gains in quality-adjusted lifetime, higher mortality reductions and lifetime gains, and fewer PSA tests needed to avoid 1 PCa death. Numbers of RP-related deaths and other RP- and RT-related adverse events tend to be lower with the original model. However, numbers of PSA tests, including false-positive PSA results, are slightly more frequent, which may be explained by the higher proportion of men remaining in the ''no cancer'' state. A large discrepancy is found for overdiagnosis, where proportions of overdiagnosis in screen-detected cancers are 40% to 77% lower than predicted by the recalibrated model, depending on screening strategy.

DISCUSSION

The objective of our work was to investigate the effect of the size of the latent prevalence pool on the benefits-harm balance of PCa screening. We recalibrated the natural history and detection component of the ONCOTYROL PCOP model adopted from the 2003 MISCAN model to match the higher PCa prevalence observed in autopsy studies.

Calibration

An important finding was that prevalence data could not be fitted by the parameters applied in the calibration of the 2003 MISCAN model. To compensate for the increase in the latent prevalence pool, additional parameters, allowing for an interruption of disease progression in the stage- and grade-specific cancer states, had to be introduced. Calibrated values for these parameters intuitively make sense as they indicate that the likelihood of cancer progression increases both with cancer stage and grade.

Increase in latent prevalence was also compensated by a decrease in stage-specific screening sensitivity. In addition, fitting of age-specific detection rates observed in the ERSPC required implementation of an effect modifier allowing for lower sensitivity of screening above age 70 years. That increasing the latent prevalence pool forces calibrated screening sensitivity to decline is plausible, since detection rates observed with screening would otherwise be overestimated. In the PCOP and MISCAN models, screening is modeled as a single test consisting of PSA testing and consecutive biopsy. As biopsy is only performed following a positive PSA result, sensitivity of screening cannot exceed the sensitivity of PSA testing. Results from pooled analyses of screening studies suggest that sensitivity of PSA testing, even with a cutoff value of 3.0 ng/mL, might be as low as 32%, with higher sensitivity of 68% for $G > 7$ tumors.[52] Compared with this, sensitivities estimated by our recalibration appear to be more in line than the higher sensitivities applied in the original model. In our calibration, we also tested the effect of grade-specific sensitivities. However, their use did not yield a better fit. In contrast, allowing sensitivity to vary with age resulted in considerable reduction of the fitting error. To our knowledge, a decrease of screening sensitivity beyond age 70 years has not been reported before. Therefore, we can only speculate about possible explanations. One hypothesis would be that older men are more likely to have consulted an urologist for various prostate problems, with whom larger or otherwise more easily detectable prostate cancers may already have been found. This selection effect (''spectrum bias'') might cause a relative enrichment of cancers that are difficult to detect by PSA screening in older men, which may explain the lower sensitivity of screening. Another hypothesis would be that older men may be more reluctant to receive a biopsy following a positive PSA test, which would reduce cancer detection by screening as well. This explanation was already proposed by Wever and others[63] to explain lower screening sensitivities in the United States compared with Europe. Other explanations might be that sensitivity of biopsy

decreases with age-related increase of prostate size or that physicians apply higher PSA cutoffs for elder men as PSA levels increase with age.

Our calibration reveals that assumptions about the size of the latent prevalence pool are crucial for screening models that use calibration to observed incidence data to simulate disease progression in the latent phase. As our recalibration showed, observed data are fitted differently depending on prevalence assumptions. Increase in prevalence was compensated by lower screening sensitivity and a slowdown of disease progression, which increased the chance to die before clinical manifestation and thereby the risk of overdiagnosis. That higher prevalence is associated with slower disease progression is plausible, because exits from the prevalence pool cannot be increased without exceeding observed incidence rates.

Impact of Latent Prevalence on the Benefit-Harm Balance of Screening

We compared 2 models that are identical except for natural history and detection parameters, which were derived by calibration under different prevalence assumptions. Both models yield contrary results concerning the benefits-harm balance of PCa screening. While the model assuming a small latent prevalence pool predicts QALY gains, the model with large prevalence pool predicts QALY losses. Comparing the outputs from both models indicates that contrary results are primarily caused by difference in overdiagnosis and screening sensitivity.

We tried to judge whether the recalibration led to clinically significant changes in the model. The answer to this question may be specific to the context and setting (e.g., risk groups v. general screening) and ultimately can be answered only when the model is filled with parameters specific to the context of interest. There is an indication, however, that changes are indeed clinically significant: with the parameters we have used for our calibration exercise, the recalibrated PCOP model still predicts gains in life expectancy with screening, but quality-adjusted life expectancy changed from gains into losses, caused by more frequent overdiagnosis and treatment-related complications. From these results, we conclude that the application of the recalibration is likely to lead to more conservative (restrictive) recommendations of screening. Whether the decision for screening will be supported remains to be seen in actual future model applications. Finally, we think that results of our recalibration are likely to initiate a new debate on the problem of overdiagnosis among both modelers and clinicians.

Which of our model predictions is more valid remains a matter of debate, not only because our analyses do not represent a comprehensive evaluation. Most other published models considering QoL[18,22,23,31,37,40] have found that benefits of screening could be outweighed by its harms, which is more consistent with results from the recalibrated PCOP model. On the other hand, results from the original PCOP model are closer to results from the 2012 MISCAN model[14] and a model published by Underwood and others.[36] Benefits of screening depend on the amount of potentially preventable mortality. Lifetime risks of PCa cancer death without screening predicted by the recalibrated and original PCOP model are 1.6% and 2.6%, respectively. That the latter exactly matches the estimate from SEER data[64] suggests that the potential benefit of PCa screening might be underestimated by the recalibrated model. Lifetime risk estimates for Austria are not available. However, the current risk of dying of PCa before age 75 years is estimated to be 0.8%.[2] It should be noted that lifetime gains predicted by our models are low compared with model predictions for breast, colorectal, and cervical cancer screening, which for screening strategies recommended by the USPSTF, for example, lie in the range of 44 to 90 days per person screened.[65–70]

Although our models apply identical utility parameters, the difference in the benefits-harm balance should also depend on the magnitude of screening- and treatment-related disutilities. Applying less severe disutilities should more strongly improve the benefit-harm balance from the recalibrated model, which predicts a higher number of cancer diagnoses, especially with frequent interval screening. Whether more favorable utility weights might reverse the negative benefit-harm balance will be investigated in a comprehensive evaluation.

Limitations

Our work has several limitations. The main limitation of both models (i.e., the original and the recalibrated PCOP model) is that they are based on calibration, which, depending on the size of the latent prevalence pool, led to contrary benefits-harm predictions. Thus, which prediction is more valid seems to depend on the validity of the prevalence assumption applied for calibration. Whereas parameters of the original PCOP model were not explicitly calibrated against observed prevalence data, which results in a small latent prevalence

pool, our recalibrated model assumes that the size of the latent prevalence pool would be best reflected by data from autopsy studies performed in the prescreening era. Whether this assumption holds may be questioned, because autopsies may detect cancers that, due to size or location, might be missed by biopsy. Beyond that, it is unknown how autopsy-detected cancers correlate with PSA increase. Questions may also concern the representativeness of the autopsy data applied in our calibration. We used age-specific latent prevalence data, which had been used by Etzioni and others,[45] to calculate the incidence of latent PCa. These data are roughly consistent with data from other studies, which rather suggest an even higher latent prevalence.[43,44,46–48]

Cure rates applied in our model are based on reductions of PCa mortality observed in the ERSPC. Strength of evidence from that trial has been questioned due to several methodological limitations.[71] Therefore, cure rates in our model might be questioned as well. However, as more recent analyses of trial data indicate that mortality reduction increases with study duration,[72] cure rates assumed by our model might still be too low, resulting in underestimated screening benefits.

Both of our models may also be biased against screening due to the use of one-time utility decrements for biopsy, prostatectomy, and radiotherapy, which appear to be rather high but were applied to maintain comparability of our results with results from the 2012 MISCAN model.[14] On the other hand, our models may be biased in favor of screening by applying a rather low estimate for perioperative mortality of prostatectomy. Both parameters should be carefully reconsidered in a comprehensive benefit-harm analysis. Again, we want to emphasize that our results should not be interpreted as results of a benefit-harm analysis but rather as a recalibration exercise. We therefore deliberately did not further discuss the transferability of our results to various national screening contexts, as we do not want to mislead the reader to interpret our findings as results of a comprehensive benefit-harm evaluation of prostate cancer screening. A comprehensive benefit-harm evaluation will be the topic of consecutive work in which we will then also address the aspect of transferability, consistency with existing guidelines, and the abovementioned magnitude and effect of treatment-induced short-term disutilities in more detail.

Furthermore, our analysis did not focus on screening of high-risk groups but on mass screening in the general population. A more ''personalized'' screening approach considering individual risks of screenees is likely to show a better benefit-harm balance for screening.

Finally, we did not factor into the model peace and anguish of mind after negative and positive screening results, although these may be important factors for individual well-being.

CONCLUSION

Calibration can help to better understand latent disease processes and to derive new hypotheses. Our calibration suggests that not all prostate cancers are at risk to progress and that the screening sensitivity may be lower at higher ages.

Our model comparison indicates that PCa screening models, which use calibration to simulate disease progression in the unobservable latent phase, are highly sensitive to assumptions about the size of the latent prevalence pool.

Recalibration of the ONCOTYROL PCOP model to higher prevalence resulted in a considerable increase in overdiagnosis and decline in screening sensitivity, which significantly worsened the benefit-harm balance of PCa screening.

As our analysis did not consider personalized screening approaches, further analyses are needed to investigate whether screening and early treatment might be beneficial for individual men with elevated PCa risk and minor weighting of treatment-related function losses. Screening combined with active surveillance, which reduces the risk of overdiagnosis, should then be considered as an alternative option as well.

ACKNOWLEDGMENT

We thank the Department of Public Health at the Erasmus Medical Center, in particular Drs Eveline A. Heijnsdijk and Gerrit Draisma, for the support and intellectual input making this study possible.

REFERENCES

1. IARC. GLOBOCAN 2012 (IARC) population fact sheet WHO EUROPE REGION (EURO). Available from: http://globocan.iarc.fr/Pages/fact_sheets_population.aspx

2. Zielonke N, Hackl M, Baldaszti R. Krebsinzidenz und Krebsmortalität in Österreich 2014. Statistik Austria, Vienna, Austria: 2014.

3. Oberaigner W, Horninger W, Klocker H, Schonitzer D, Stuhlinger W, Bartsch G. Reduction of prostate cancer mortality in Tyrol,

Austria, after introduction of prostate-specific antigen testing. Am J Epidemiol. 2006;164(4):376–84.

4. Oberaigner W, Siebert U, Horninger W, et al. Prostate-specific antigen testing in Tyrol, Austria: prostate cancer mortality reduction was supported by an update with mortality data up to 2008. Int J Public Health. 2012;57(1):57–62.

5. Horninger W, Berger A, Pelzer A, et al. Screening for prostate cancer: updated experience from the Tyrol study. Can J Urol. 2005;12(Suppl 1):7–13; discussion 92–3.

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

7. Schroder FH, Hugosson J, Carlsson S, et al. Screening for prostate cancer decreases the risk of developing metastatic disease: findings from the European Randomized Study of Screening for Prostate Cancer (ERSPC). Eur Urol. 2012;62(5):745–52.

8. Andriole GL, Crawford ED, Grubb RL III, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med. 2009;360(13):1310–9.

9. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. Cochrane Database Syst Rev. 2013;1:CD004720.

10. Chou R, Croswell JM, Dana T, et al. Screening for prostate cancer: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2011;155(11):762–71.

11. Chou R, Dana T, Bougatsos C, et al. Treatments for localized prostate cancer: systematic review to update the 2002 U.S. Preventive Services Task Force Recommendation [Internet]. 2011 Oct U.S. Preventive Services Task Force Recommendation. Available from: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0032891/

12. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst. 2003;95(12):868–78.

13. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst. 2009;101(6):374–83.

14. Heijnsdijk EA, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. N Engl J Med. 2012; 367(7):595–605.

15. Heidenreich A, Bastian PJ, Bellmunt J, et al. EAU guidelines on prostate cancer, part 1: screening, diagnosis, and local treatment with curative intent—update 2013. Eur Urol. 2014;65(1):124–37.

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

17. Optenberg SA, Thompson IM. Economics of screening for carcinoma of the prostate. Urol Clin North Am. 1990;17(4):719–37.

18. Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS. Screening for prostate cancer: a decision analytic view. JAMA. 1994;272(10):773–80.

19. Littrup PJ, Kane RA, Mettlin CJ, et al. Cost-effective prostate cancer detection: reduction of low-yield biopsies. Investigators of the American Cancer Society National Prostate Cancer Detection Project. Cancer. 1994;74(12):3146–58.

20. Barry MJ, Fleming C, Coley CM, Wasson JH, Fahs MC, Oesterling JE. Should Medicare provide reimbursement for prostate-specific antigen testing for early detection of prostate cancer? Part IV: Estimating the risks and benefits of an early detection program. Urology. 1995;46(4):445–61.

21. Gustafsson O, Carlsson P, Norming U, Nyman CR, Svensson H. Cost-effectiveness analysis in early detection of prostate cancer: an evaluation of six screening strategies in a randomly selected population of 2,400 men. Prostate. 1995;26(6):299–309.

22. Cantor SB, Spann SJ, Volk RJ, Cardenas MP, Warren MM. Prostate cancer screening: a decision analysis. J Fam Pract. 1995;41(1): 33–41.

23. Gottlieb RH, Mooney C, Mushlin AI, Rubens DJ, Fultz PJ. The prostate: decreasing cost-effectiveness of biopsy with advancing age. Invest Radiol. 1996;31(2):84–90.

24. Coley CM, Barry MJ, Fleming C, Fahs MC, Mulley AG. Early detection of prostate cancer, part II: estimating the risks, benefits, and costs. American College of Physicians. Ann Intern Med. 1997;126(6):468–79.

25. Perez-Niddam K, Thoral F, Charvet-Protat S. Economic evaluation of a prostate cancer screening program in France: a decision model. Crit Rev Oncol Hematol. 1999;32(2):167–73.

26. Etzioni R, Cha R, Cowen ME. Serial prostate specific antigen screening for prostate cancer: a computer model evaluates competing strategies. J Urol. 1999;162(3, Pt 1):741–8.

27. Ross KS, Carter HB, Pearson JD, Guess HA. Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. JAMA. 2000;284(11):1399–405.

28. Benoit RM, Gronberg H, Naslund MJ. A quantitative analysis of the costs and benefits of prostate cancer screening. Prostate Cancer Prostatic Dis. 2001;4(3):138–45.

29. Ellison L, Cheli CD, Bright S, Veltri RW, Partin AW. Cost-benefit analysis of total, free/total, and complexed prostate-specific antigen for prostate cancer screening. Urology. 2002;60(4, Suppl 1):42–6.

30. Sennfalt K, Sandblom G, Carlsson P, Varenhorst E. Costs and effects of prostate cancer screening in Sweden—a 15-year follow-up of a randomized trial. Scand J Urol Nephrol. 2004; 38(4):291–8.

31. Howard DH. Life expectancy and the value of early detection. J Health Econ. 2005;24(5):891–906.

32. Tsodikov A, Szabo A, Wegelin J. A population model of prostate cancer incidence. Stat Med. 2006;25(16):2846–66.

33. Kobayashi T, Goto R, Ito K, Mitsumori K. Prostate cancer screening strategies with re-screening interval determined by individual baseline prostate-specific antigen values are cost-effective. Eur J Surg Oncol. 2007;33(6):783–9.

34. Howard K, Barratt A, Mann GJ, Patel MI. A model of prostate-specific antigen screening outcomes for low- to high-risk men: information to support informed choices. Arch Intern Med. 2009; 169(17):1603–10.

35. Shteynshlyuger A, Andriole GL. Cost-effectiveness of prostate specific antigen screening in the United States: extrapolating from the European study of screening for prostate cancer. J Urol. 2011; 185(3):828–32.

36. Underwood DJ, Zhang J, Denton BT, Shah ND, Inman BA. Simulation optimization of PSA-threshold based prostate cancer screening policies. Health Care Manag Sci. 2012;15(4):293–309.

37. Zhang J, Denton BT, Balasubramanian H, Shah ND, Inman BA. Optimization of PSA screening policies: a comparison of the patient and societal perspectives. Med Decis Making. 2012;32(2):337–49.

38. Wu GH, Auvinen A, Yen AM, Hakama M, Walter SD, Chen HH. A stochastic model for survival of early prostate cancer with adjustments for leadtime, length bias, and over-detection. Biom J. 2012; 54(1):20–44.

39. Gulati R, Gore JL, Etzioni R. Comparative effectiveness of alternative prostate-specific antigen-based prostate cancer screening strategies: model estimates of potential benefits and harms. Ann Intern Med. 2013;158(3):145–53.

40. Martin AJ, Lord SJ, Verry HE, Stockler MR, Emery JD. Risk assessment to guide prostate cancer screening decisions: a cost-effectiveness analysis. Med J Aust. 2013;198(10):546–50.

41. Heijnsdijk EA, der Kinderen A, Wever EM, Draisma G, Roobol MJ, de Koning HJ. Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer. Br J Cancer. 2009;101(11):1833–8.

42. Wever EM, Hugosson J, Heijnsdijk EA, Bangma CH, Draisma G, de Koning HJ. To be screened or not to be screened? Modeling the consequences of PSA screening for the individual. Br J Cancer. 2012;107(5):778–84.

43. Breslow N, Chan CW, Dhom G, et al. Latent carcinoma of prostate at autopsy in seven areas. The International Agency for Research on Cancer, Lyons, France. Int J Cancer. 1977;20(5):680–8.

44. Sakr WA, Grignon DJ, Haas GP, Heilbrun LK, Pontes JE, Crissman JD. Age and racial distribution of prostatic intraepithelial neoplasia. Eur Urol. 1996;30(2):138–44.

45. Etzioni R, Cha R, Feuer EJ, Davidov O. Asymptomatic incidence and duration of prostate cancer. Am J Epidemiol. 1998; 148(8):775–85.

46. Sanchez-Chapado M, Olmedilla G, Cabeza M, Donat E, Ruiz A. Prevalence of prostate cancer and prostatic intraepithelial neoplasia in Caucasian Mediterranean males: an autopsy study. Prostate. 2003;54(3):238–47.

47. Rich AR. On the frequency of occurrence of occult carcinoma of the prostrate. 1934. Int J Epidemiol. 2007;36(2):274–7.

48. Haas GP, Delongchamps NB, Jones RF, et al. Needle biopsies on autopsy prostates: sensitivity of cancer detection based on true prevalence. J Natl Cancer Inst. 2007;99(19):1484–9.

49. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Med Decis Making. 2012;32(5):690–700.

50. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making. 2012;32(5):722–32.

51. Eurostat. European Union Open Data Portal, life table. Available from: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_mlifetable&lang=en

52. 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(2):70–98.

53. Fink KG, Hutarew G, Lumper W, Jungwirth A, Dietze O, Schmeller NT. Prostate cancer detection with two sets of ten-core compared with two sets of sextant biopsies. Urology. 2001;58(5):735–9.

54. Shariat SF, Roehrborn CG. Using biopsy to detect prostate cancer. Rev Urol. 2008;10(4):262–80.

55. National Cancer Institute. SEER Cancer Statistics Review, 1975–2011. Available from: http://seer.cancer.gov/csr/1975_2011/

56. Bartsch G, Horninger W, Klocker H, et al. Tyrol Prostate Cancer Demonstration Project: early detection, treatment, outcome, incidence and mortality. BJU Int. 2008;101(7):809–16.

57. Wever EM, Draisma G, Heijnsdijk EA, de Koning HJ. How does early detection by screening affect disease progression? Modeling estimated benefits in prostate cancer screening. Med Decis Making. 2011;31(4):550–8.

58. Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med. 2009;360(13):1320–8.

59. Korfage IJ, Essink-Bot ML, Borsboom GJ, et al. Five-year follow-up of health-related quality of life after primary treatment of localized prostate cancer. Int J Cancer. 2005;116(2):291–6.

60. Olsen J, Jepsen MR. Human papillomavirus transmission and cost-effectiveness of introducing quadrivalent HPV vaccination in Denmark. Int J Technol Assess Health Care. 2010;26(2):183–91.

61. Krahn MD, Bremner KE, Alibhai SM, et al. A reference set of health utilities for long-term survivors of prostate cancer: population-based data from Ontario, Canada. Qual Life Res. 2013;22(10):2951–62.

62. Statistik Austria. Sterbetafel 2010/2012 männlich. Available from: https://www.statistik.at/web_de/static/sterbetafeln_186871_bis_201012_nach_dem_geschlecht_022541.xlsx

63. Wever EM, Draisma G, Heijnsdijk EA, et al. Prostate-specific antigen screening in the United States vs in the European Randomized Study of Screening for Prostate Cancer–Rotterdam. J Natl Cancer Inst. 2010;102(5):352–5.

64. National Cancer Institute. SEER Cancer Statistics Review, 1975–2008. Available from: http://seer.cancer.gov/csr/1975_2008/

65. US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009;151(10):716–26, W-236.

66. 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(10):738–47.

67. US Preventive Services Task Force. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2008;149(9):627–37.

68. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med. 2008;149(9):659–69.

69. Moyer VA. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012; 156(12):880–91, W312.

70. Kulasingam SL, Havrilesky L, Ghebre R, Myers ER. Screening for Cervical Cancer: A Decision Analysis for the U.S. Preventive Services Task Force. 2011 May; Agency for Healthcare Research and Quality, Rockville (MD), USA, Report No. 11-05157-EF-1.

71. Barry MJ. Screening for prostate cancer—the controversy that refuses to die. N Engl J Med. 2009;360(13):1351–4.

72. Schroder 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(9959):2027–35.