Personalized risks of overdiagnosis for screen-detected prostate cancer incorporating patient comorbidities: Estimation and communication

Roman Gulati, **Sarah P. Psutka**, **Ruth Etzioni**

Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington (RG, RE) and Department of Urology, University of Washington, Seattle, Washington (SPP)

Abstract

Purpose: Shared patient-physician decision-making regarding treatment for prostate cancer detected by prostate-specific antigen screening involves a complex calculus weighing the risk of the cancer and patient life expectancy. We investigated quantifying these competing risks using the probability that the cancer was "overdiagnosed"—i.e., would not have been clinically diagnosed (diagnosed without screening) during the patient's remaining lifetime.

Materials and Methods: Using an established model of prostate cancer screening and clinical diagnosis, we simulated screen-detected cases and determined whether modeled clinical diagnosis would occur before non-cancer death, which was based on comorbidity-adjusted population lifetables. Logistic regression models were fitted to the simulated data and used to estimate overdiagnosis probabilities given patient age, PSA level, Gleason sum, and comorbidity category. An online calculator was developed to communicate overdiagnosis estimates; face validity and ease of use was assessed by surveying 32 clinical experts.

Results: Estimated probabilities of overdiagnosis ranged 4%–78% across clinicopathologic variables and comorbidity status. Ignoring comorbidity, the estimated probability for a 70-year-old man with PSA 9.4 ng/mL and Gleason 6 is 34%; if he has severe comorbidities, the estimate increases to 51%, a personalization that may help inform the choice between active surveillance and definitive treatment. Based on responses from 20/32 experts, we modified the online calculator's explanation of overdiagnosis and input method for comorbid conditions.

Conclusions: The probability of overdiagnosis is strongly influenced by comorbidity status in addition to age. Personalized estimates incorporating comorbidity may contribute to shared decision-making between patients and providers regarding personalized treatment selection.

Keywords

comorbidity; mass screening; medical overuse; prostatic neoplasms

Correspondence: Roman Gulati, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-B230, Seattle, WA 98109-1024. Tel: +1 206 6677795. Fax: +1 206 6677264., rgulati@fredhutch.org.

Conflict of interest: None.

INTRODUCTION

Prostate cancer (PCa) is the most commonly diagnosed cancer and the second most common cause of cancer death among men in the United States.[1] Despite high-quality evidence that prostate-specific antigen (PSA) screening reduces PCa mortality, the United States Preventive Services Task Force first recommended against PSA screening and, more recently, recommended shared decision-making between patient and provider regarding PSA screening.[2] Appreciation of the harms of PSA screening has driven the growing acceptance of active surveillance for the management of low-to-intermediate-risk PCa.[3, 4]

Overdiagnosis occurs when a cancer detected by screening would not have been diagnosed without screening during the patient's remaining lifetime.[5] This "epidemiological definition"[6] of overdiagnosis implies that the likelihood a patient is overdiagnosed depends on the competing risks of a counterfactual future clinical diagnosis and non-cancer death (Figure 1). This is distinct from the "clinical definition," which defines overdiagnosis as detection of low-grade or low-risk PCa irrespective of the risk of non-cancer death. The two definitions are related because less aggressive cancers tend to take a longer time to progress to clinical diagnosis (thus are more likely to be overdiagnosed according to either definition). They are not the same, however, because older men or those with significant comorbidity are likely to die of non-cancer causes sooner, which increases their risk of overdiagnosis according to the epidemiological definition but not the clinical definition. In the rest of this manuscript we use the epidemiological definition of overdiagnosis. We previously published[7] estimated risks of overdiagnosis by age, PSA, and grade but did not account for comorbidity status.

Overtreatment occurs when an overdiagnosed cancer is treated with curative intent. Because, by definition, an overdiagnosed cancer would not cause morbidity or mortality in the patient's remaining lifetime, curative treatment cannot improve outcomes and can only cause harm. In addition to creating anxiety and financial hardship, treatments such as radical prostatectomy and radiotherapy are associated with significant risks of post-treatment impotence and incontinence, which can be lifelong and are associated with declines in health-related quality of life.[8] For men with a high probability of overdiagnosis, it is increasingly acknowledged that active surveillance (AS) may be more appropriate. For any individual patient, therefore, the chance that his cancer has been overdiagnosed is a potentially useful piece of information that can be factored into his decision regarding definitive treatment or AS.

The primary objective of this study is to derive personalized estimates of the risk of overdiagnosis for men with screen-detected PCa incorporating comorbidities in addition to tumor characteristics and patient age. We offer the resulting estimates as an aid to inform patient-physician decisions regarding individualized treatment choices. To provide access to these personalized estimates of overdiagnosis, we develop an online calculator and present results from a survey of clinical PCa experts regarding its face validity and usability.

MATERIALS AND METHODS

Overview

Our work leverages a previously developed microsimulation model of PCa onset, progression, and clinical diagnosis.[9, 10] Using this model, we simulate a virtual population of life histories that include age at non-cancer death from population life tables. We then superimpose PSA screening with a specified schedule and biopsy-referral criteria on these life histories, resulting in screen-detected PCa for some men. For each screen-detected patient, we compare projected times to clinical diagnosis and non-cancer death to determine whether his PCa was overdiagnosed. We then use logistic regression to model overdiagnosis as a function of patient age, disease characteristics (PSA, grade), and comorbidity category. For specified values of these variables, the fitted regression model yields estimates of individualized overdiagnosis risks, which are provided via a user-friendly online calculator.

Time to clinical diagnosis

Time to clinical diagnosis for a screen-detected PCa was projected using our microsimulation model (Figure 1),[9, 10] which generates individual PSA trajectories and ages at disease onset and clinical diagnosis. The model parameters were previously estimated to match PCa incidence trends by age, stage, and grade from the Surveillance, Epidemiology, and End Results (SEER) program under a reconstruction of population PSA testing starting in 1987.[11] The model calibrates well to observed trends in PCa incidence before and after the introduction of PSA screening (Supplemental Figure 1). Using this model, "clinical diagnosis" for a screen-detected patient refers to counterfactual future diagnosis based on clinical practice in the U.S. population before PSA screening was introduced.

Time to non-cancer death

To generate time to non-cancer death for men with life-limiting comorbid conditions, we used previously estimated comorbidity-adjusted life tables.[12, 13] Briefly, 16 comorbid conditions in the Charlson comorbidity index[14] were identified using medical claims from the linked SEER-Medicare database in the year preceding PCa diagnosis. Survival curves were estimated conditional on age and comorbidity category as defined in Table 1. Survival beyond the available follow-up smoothly reverts to survival for men at average risk of non-cancer death, tempering the effects of comorbidity on long-term non-cancer survival.[15] Example comorbidity-specific survival curves for men who were 70 years old at diagnosis are presented in Supplemental Figure 2. For comparison, we also generated time to non-cancer death for men at average risk of non-cancer death using population life tables without adjustment for comorbidity.[16]

Estimating overdiagnosis

To estimate individualized overdiagnosis risks, we simulated four cohorts of 10 million life histories each, with age at non-cancer death drawn from comorbidity-adjusted survival curves corresponding to the four comorbidity categories. We then superimposed biennial PSA screening starting at age 50, with 40% of men with PSA > 4.0 ng/mL receiving biopsy depending on age and PSA level,[17] and biopsy having 80% sensitivity to detect preclinical

PCa.[18–20] For each screen-detected PCa, we determined whether clinical diagnosis would occur before non-cancer death. The resulting binary (yes/no) indicator became the response in a logistic regression model with four covariates: patient age, PSA level, Gleason sum, and comorbidity category. For the comorbidity-independent analysis, a single cohort of 10 million life histories was simulated using population life tables without adjustment for comorbidity, and the corresponding logistic regression involved only the first three covariates. To account for possible cross-dependence on the clinicopathologic variables, all two-way interactions were used in the logistic regression models.

Discrimination (separation of overdiagnosed vs non-overdiagnosed cancer) was visualized using receiver operating characteristic (ROC) curves and quantified using areas under the ROC curves (AUCs). Calibration (reliability of predictions) was visualized using calibration plots of observed vs predicted risks based on each percentile of the observed risk.

Communicating overdiagnosis

We developed an online calculator with the following inputs: patient age at screen detection (continuous, range: 50–74 years), PSA level (continuous, range: 4.0–10.0 ng/mL), Gleason sum (2–6 or 7), and, for patients ages 66–74 at screen detection, specific life-limiting comorbid conditions. Real-time reactivity updates the estimated probability of overdiagnosis and provides a pictograph (a color-coded grid of person-shaped icons) to visualize the estimated frequency of overdiagnosis per 100 similar patients.[21]

A one-page questionnaire (Supplemental Figure 3), adapted from the Human Computer Trust rating scale,[22] was then emailed to 32 clinical PCa experts. The questionnaire included a link to the online calculator, a hypothetical example, propositions about face validity and usability of the calculator with level of agreement or disagreement to be indicated using a 5-point Likert scale, and a free-text box for comments and suggestions.

RESULTS

Accuracy of overdiagnosis estimates

The fitted logistic regression models are reported in Tables 2 and 3. As expected, the odds of overdiagnosis increased with age and comorbidity category and decreased with PSA level and Gleason sum. Discrimination and calibration of the models are shown in Figure 3. The AUC for the comorbidity-dependent model (ages 66–74) was 0.65 (95% confidence interval [CI] 0.65 to 0.66), a non-trivial improvement compared to the AUC from the same model when comorbidity is excluded (0.62; 95% CI 0.62 to 0.63). The AUC for the comorbidity-independent model (ages 50–74) was 0.70 (95% CI 0.69 to 0.70). Calibration intercepts and slopes for both models are 0.00 (95% CI –0.01 to 0.01) and 1.00 (95% CI 0.99 to 1.01), close to the theoretical ideal values of 0 and 1. The 95% CIs were very narrow due to the large number of simulated screen-detected PCa cases and do not account for uncertainty in the natural history model specification.

Illustration of overdiagnosis estimates

As shown in Figure 4, estimated probabilities of overdiagnosis vary widely across patient and tumor features, ranging from 13% to 78% in the comorbidity-dependent model and from 4% to 69% in the comorbidity-independent model.

To illustrate, the estimated probability of overdiagnosis for a 70-year-old man with PSA of 9.4 ng/mL and biopsy Gleason 6 is 34% if we ignore any health conditions. This probability decreases to 27% if he has no comorbid conditions and increases to 51% if he has diabetes and cardiovascular disease. As this example shows, accounting for patient health conditions may provide materially different direction in terms of whether AS is likely to be appropriate.

Assessment of online calculator

We received responses from 20 out of 32 experts surveyed, including 18 as completed questionnaires (Supplemental Figure 4) and another 2 as summative comments only. The experts favorably reviewed the interface usability and graphical design. However, there were mixed responses about the trust in the estimates and the expectation that the calculator would be useful in practice, particularly if used with a patient.

Based on the comments, the mistrust appeared to have been related to conflation of the clinical and epidemiological definitions of overdiagnosis.[6] According to the clinical definition, overdiagnosis equates to diagnosis of low-grade or low-risk PCa regardless of the risk of non-cancer death. In contrast, according to the epidemiological definition, grade is moderately prognostic for clinical diagnosis and not at all prognostic for non-cancer death; therefore, grade is only modestly prognostic for overdiagnosis. Other comments included recommending that we more clearly define the term "overdiagnosis"; restructure the selection of comorbidity inputs; more clearly indicate that comorbidity-specific results are only available for ages 66–74 years; make the text preamble more accessible to patients; give an example with interpretation; clarify that results are for average-risk patients with limited applicability to men with risk factors (e.g., based on race or family history); and add a descriptive legend to the pictograph. We modified the calculator taking all comments into consideration (Supplemental Figure 5). The modified calculator is available at https://rgulati.shinyapps.io/personalized-overdiagnosis-prostate.

DISCUSSION

Both the U.S. Preventive Services Task Force and the American Cancer Society recently identified the development of new tools to reduce PCa overtreatment as an important research priority.[2, 23] To the best of our knowledge, only two previous studies derived personalized estimates of the risk of overdiagnosis that could be used by patients and their providers to make informed decisions about appropriate treatment for screen-detected PCa. One used the same PCa natural history model[7]; the other compared PCa incidence in screened and unscreened groups by age and PSA level.[24] However, no studies to date have accounted for the impact of longevity-limiting comorbidities on the estimated risk of overdiagnosis.

In the current study, we derived personalized estimates of the risk of overdiagnosis using patient age, PSA level, Gleason sum, and burden of comorbid conditions. Using this framework, we observed that the estimated probabilities of overdiagnosis vary widely, for example from 13% for men with no comorbid conditions at age 66 with PSA of 10.0 ng/mL and Gleason sum 7 up to 78% for men with severe comorbid conditions at age 74 with PSA of 4.0 ng/mL and Gleason sum 2–6. These estimates represent a meaningful measure to weigh the competing risks of PCa against the baseline comorbidity burden and may help further inform decision-making regarding the relative risks and benefits of definitive treatment vs active surveillance. For example, men with lower estimated probabilities of overdiagnosis may be considered good candidates for definitive treatment, while men with higher estimated probabilities may be better served by initiating active surveillance or in extreme cases followed without curative intent. In general, the estimated risks of overdiagnosis are intended to be used in conjunction with other risk assessment tools, and final treatment decisions should account for patient preferences.

A prototype online calculator was favorably rated by a panel of clinical PCa experts in terms of usability. The calculator was modified in response to comments regarding confusion about the definition of overdiagnosis and related terminology as it relates to this study, the input method for patient comorbid conditions, and the inclusion of additional explanatory information.

This study leverages a model of PCa natural history that was rigorously estimated using population-based data sources on serum PSA levels, historical PSA screening rates, and PCa incidence rates in the U.S. Because the model represents the disease process with and without screening at the individual level, estimated risks of overdiagnosis can be based upon individual clinicopathologic variables in addition to age. The absolute and relative performance of these risk estimates depend on the adequacy of the underlying model of PCa natural history and the modeled dependence of overdiagnosis on patient age, PSA level, Gleason sum, and comorbid conditions.

The main limitation of this study is that the underlying model is an approximation of complex biological processes. While the base model has previously been shown to reasonably reproduce trends in PCa incidence by age, stage, and grade in the U.S.[25] and in randomized PCa screening trials,[26] estimated frequencies of overdiagnosis are not observable and therefore cannot be empirically validated. A secondary limitation is that contemporary screening and diagnostic practices, which can involve medical imaging and/or additional biomarkers, could not be considered because population-based longitudinal data on how these measurements correlate with PSA concentrations and PCa natural history are not available. Further, because the natural history model was calibrated to historical SEER incidence data, it may not reflect contemporary grading procedures, which are based on more extensive biopsies, newer imaging modalities, and evolving pathology standards.

Given these limitations, our predictions are most relevant for average-risk, asymptomatic men who were screened approximately biennially after age 50 years, who had PSA concentration between 4 and 10 ng/mL, and who elected biopsy without further diagnostic evaluation. To the extent that additional factors contributed to decisions to receive a PSA test

or biopsy, and to the extent that these factors are associated with higher PCa prevalence and/or increased aggressiveness, the predicted risk of overdiagnosis may be overestimated.

Future extensions of this work could utilize comorbidity-adjusted life tables for a wider age range and evaluate whether the calculator has a positive impact on the patient-physician decision-making process about appropriate treatment.

CONCLUSIONS

We find that the probability of overdiagnosis varies widely across patients when incorporating age and comorbidity. We present an online tool that provides estimated probabilities of overdiagnosis accounting for these individual patient factors. These personalized estimates of the probability that a screen-detected PCa was overdiagnosed, and therefore would not cause symptoms during the patient's lifetime or threaten his longevity, may help to inform shared decision-making regarding treatment options for newly diagnosed disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

References

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2018. CA Cancer J Clin, 68: 7, 2018 [PubMed: 29313949]

2. Grossman DC, Curry SJ, Owens DK et al.: Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. JAMA, 319: 1901, 2018 [PubMed: 29801017]

3. Perlis N, Klotz L: Contemporary Active Surveillance: Candidate Selection, Follow-up Tools, and Expected Outcomes. Urol Clin North Am, 44: 565, 2017 [PubMed: 29107273]

4. Carroll PH, Mohler JL: NCCN Guidelines Updates: Prostate Cancer and Prostate Cancer Early Detection. J Natl Compr Canc Netw, 16: 620, 2018 [PubMed: 29784740]

5. Davies L, Petitti DB, Woo M et al.: Defining, Estimating, and Communicating Overdiagnosis in Cancer Screening. Ann Intern Med, 169: 739, 2018

6. Loeb S, Bjurlin MA, Nicholson J et al.: Overdiagnosis and overtreatment of prostate cancer. Eur Urol, 65: 1046, 2014 [PubMed: 24439788]

7. Gulati R, Inoue LY, Gore JL et al.: Individualized Estimates of Overdiagnosis in Screen-Detected Prostate Cancer. J Natl Cancer Inst, 106: djt367, 2014 [PubMed: 24399850]

8. Donovan JL, Hamdy FC, Lane JA et al.: Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med, 375: 1425, 2016 [PubMed: 27626365]

9. Gulati R, Inoue L, Katcher J et al.: Calibrating disease progression models using population data: a critical precursor to policy development in cancer control. Biostatistics, 11: 707, 2010 [PubMed: 20530126]

10. 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, 158: 145, 2013 [PubMed: 23381039]

11. Mariotto AB, Etzioni R, Krapcho M et al.: Reconstructing PSA testing patterns between black and white men in the US from Medicare claims and the National Health Interview Survey. Cancer, 109: 1877, 2007 [PubMed: 17372918]

12. Cho H, Klabunde CN, Yabroff KR et al.: Comorbidity-adjusted life expectancy: a new tool to inform recommendations for optimal screening strategies. Ann Intern Med, 159: 667, 2013 [PubMed: 24247672]

13. Cho H, Mariotto A, Mann BS et al.: Assessing Non-Cancer-Related Health Status of US Cancer Patients: Other-Cause Survival and Comorbidity Prevalence. Am J Epidemiol, 2013

14. Charlson ME, Pompei P, Ales KL et al.: A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis, 40: 373, 1987 [PubMed: 3558716]

15. Lansdorp-Vogelaar I, Gulati R, Mariotto AB et al.: Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. Ann Intern Med, 161: 104, 2014 [PubMed: 25023249]

16. National Center for Health Statistics: Vital statistics of the United States, Volume II: Mortality, part A Washington DC: Government Printing Office, various years

17. Grubb RL 3rd, Pinsky PF, Greenlee RT et al.: Prostate cancer screening in the Prostate, Lung, Colorectal and Ovarian cancer screening trial: update on findings from the initial four rounds of screening in a randomized trial. BJU Int, 102: 1524, 2008 [PubMed: 19035857]

18. Babaian RJ, Toi A, Kamoi K et al.: A comparative analysis of sextant and an extended 11-core multisite directed biopsy strategy. J Urol, 163: 152, 2000 [PubMed: 10604335]

19. Fink KG, Hutarew G, Lumper W et al.: Prostate cancer detection with two sets of ten-core compared with two sets of sextant biopsies. Urology, 58: 735, 2001 [PubMed: 11711351]

20. Serefoglu EC, Altinova S, Ugras NS et al.: How reliable is 12-core prostate biopsy procedure in the detection of prostate cancer? Can Urol Assoc J, 7: E293, 2013 [PubMed: 22398204]

21. Lipkus IM: Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. Med Decis Making, 27: 696, 2007 [PubMed: 17873259]

22. Madsen M, Gregor S: Measuring human-computer trust. Presented at the Eleventh Australasian Conference on Information Systems, Brisbane, 6–8 12, 2000

23. Wender RC, Brawley OW, Fedewa SA et al.: A blueprint for cancer screening and early detection: Advancing screening's contribution to cancer control. CA Cancer J Clin, 2018

24. Vickers AJ, Sjoberg DD, Ulmert D et al.: Empirical estimates of prostate cancer overdiagnosis by age and prostate-specific antigen. BMC Med, 12: 26, 2014 [PubMed: 24512643]

25. Gulati R, Tsodikov A, Etzioni R et al.: Expected population impacts of discontinued prostate-specific antigen screening. Cancer, 120: 3519, 2014 [PubMed: 25065910]

26. De Koning HJ, Gulati R, Moss SM et al.: The efficacy of PSA screening: impact of key components in the ERSPC and PLCO trial. Cancer, 124: 1197, 2018 [PubMed: 29211316]
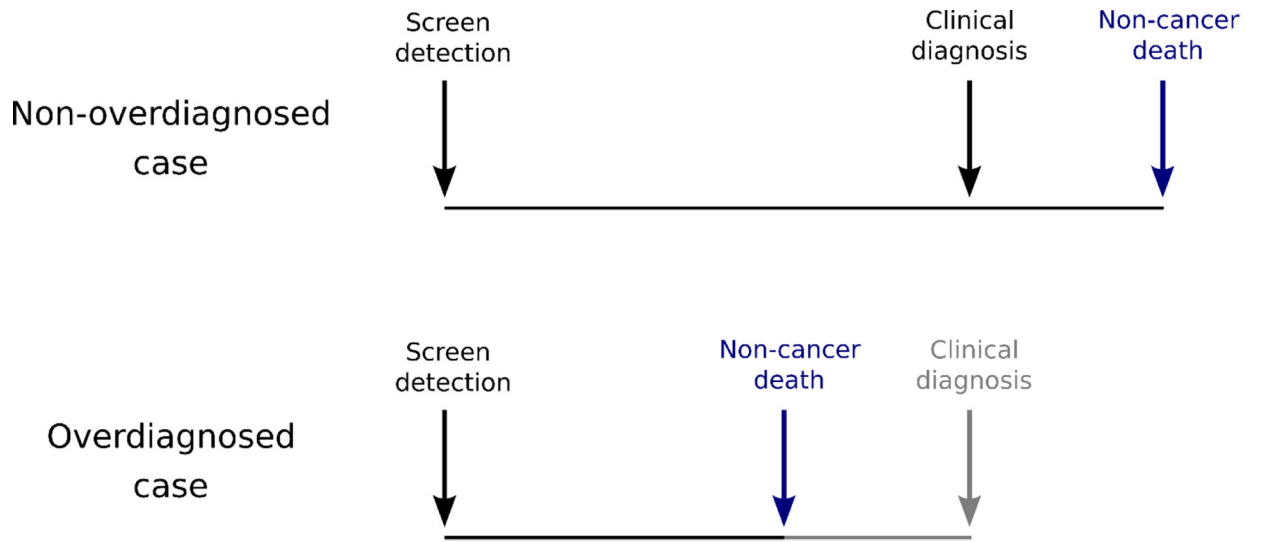
Figure 1.
Diagram illustrating how overdiagnosis of a cancer detected by screening is determined by the competition between clinical diagnosis and non-cancer death.
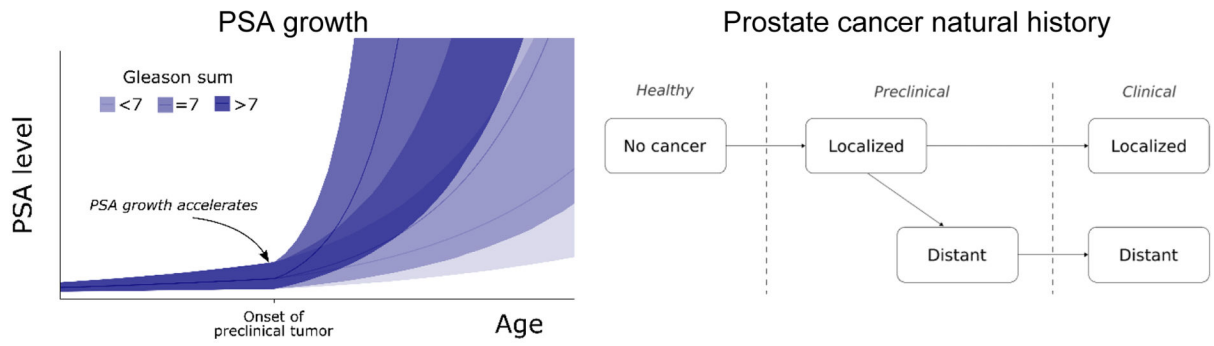
Figure 2.
Diagram of linked model of PSA growth prostate cancer natural history used to generate clinical diagnosis for simulated prostate cancers detected by screening.
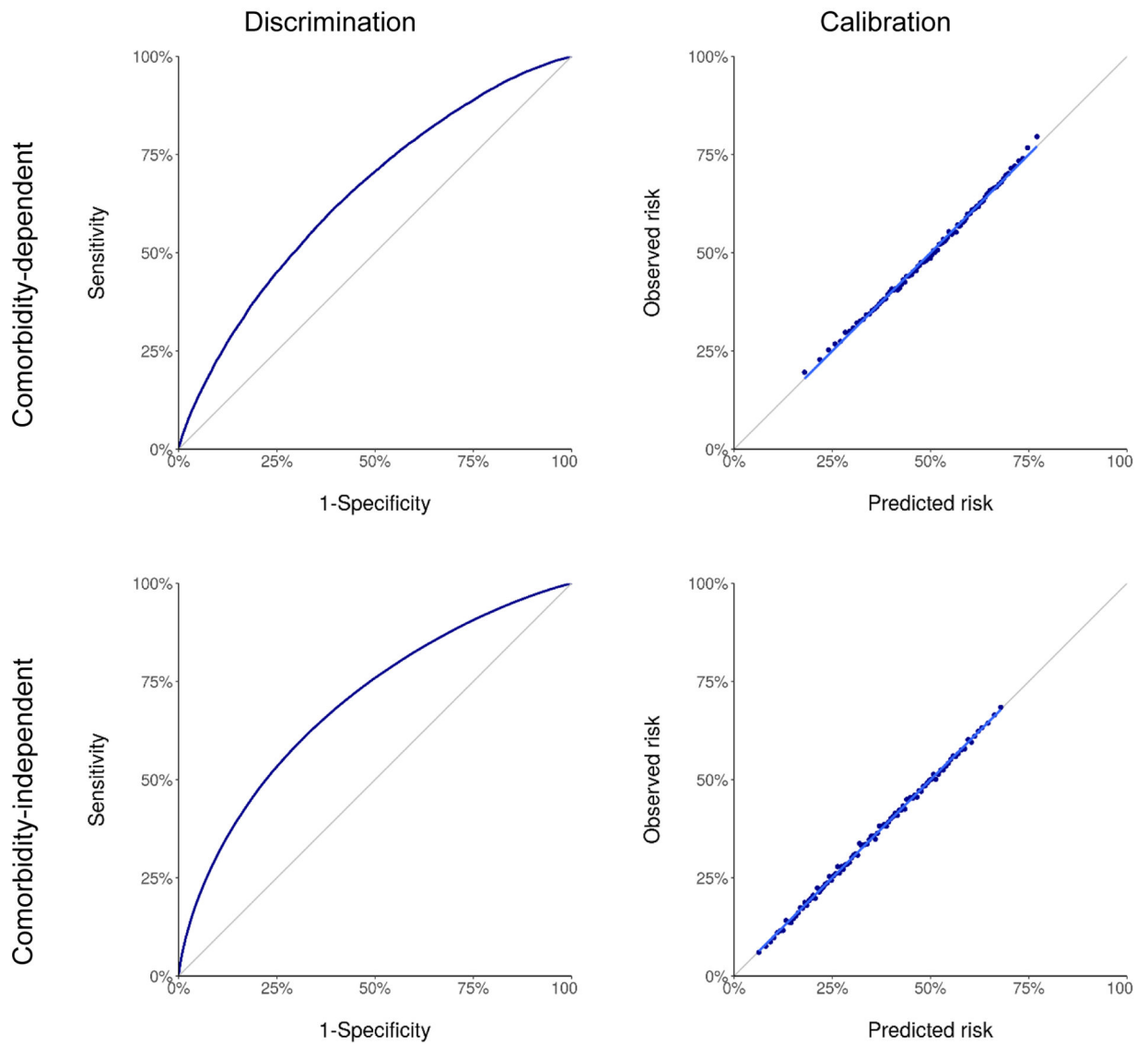
Discrimination

Calibration

Figure 3.
Performance of estimated probabilities of prostate cancer overdiagnosis based on comorbidity-dependent (ages 66–74 years) and comorbidity-independent (ages 50–74 years) logistic regression models.
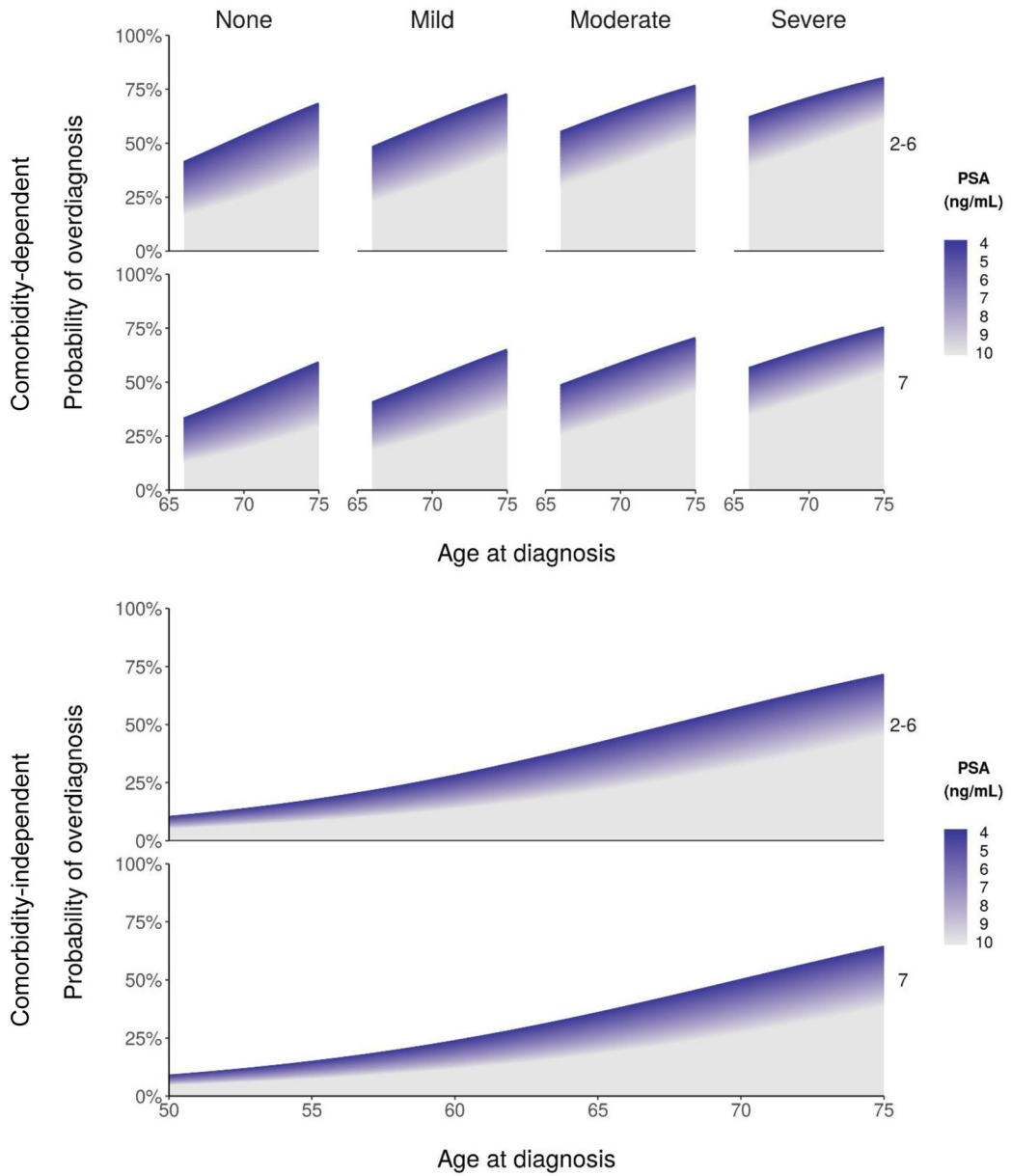
Figure 4.
Estimated probabilities of prostate cancer overdiagnosis for patients based on comorbidity-dependent (ages 66–74 years) and comorbidity-independent (ages 50–74 years) logistic regression models.

Note: While some patients with PSA > 10 ng/mL may be overdiagnosed, this probability was not estimated in this study.

Table 1.

Definitions of comorbidity groups

Comorbidity group	Comorbid conditions
None	None
Mild	History of myocardial infarction, acute myocardial infarction, ulcer, or rheumatologic disease
Moderate	Cardiovascular disease; paralysis; diabetes; or combinations of diabetes with myocardial infarction, ulcer, or rheumatologic disease
Severe	Acquired immune deficiency syndrome, chronic obstructive pulmonary disease, mild or severe liver disease, chronic renal failure, dementia congestive heart failure, or combinations of conditions not listed above

Table 2.

Fitted logistic regression model of prostate cancer overdiagnosis for patients with specific comorbid conditions, ages 66–74 years

Covariate	Odds ratio	95% CI	P value
Intercept[*]	0.53	(0.52, 0.54)	<0.001
Age at diagnosis, y	1.14	(1.14, 1.15)	<0.001
PSA level, ng/mL	0.80	(0.79, 0.80)	<0.001
Gleason 7 vs 2–6	0.68	(0.66, 0.70)	<0.001
Comorbidity[†]	1.33	(1.32, 1.33)	<0.001
Age × PSA	1.00	(1.00, 1.00)	0.11
Age × (Gleason 7 vs 2–6)	0.99	(0.99, 1.00)	<0.001
Age × Comorbidity[†]	0.99	(0.99, 0.99)	<0.001
PSA × (Gleason 7 vs 2–6)	1.01	(1.00, 1.01)	0.04
PSA × Comorbidity[†]	1.02	(1.02, 1.02)	<0.001
(Gleason 7 vs 2–6) × Comorbidity[†]	1.04	(1.03, 1.05)	<0.001

[*] Covariates were scaled so the intercept corresponds to a patient age 66 y with PSA 4 ng/mL and Gleason 2–6 at screen detection.

[†] Comorbidity is a continuous coding of comorbidity groups none, mild, moderate, and severe.

Table 3.

Fitted logistic regression model of prostate cancer overdiagnosis for patients with average comorbidities, ages 50–74 years

Covariate	Odds ratio	95% CI	P value
Intercept[*]	0.11	(0.11, 0.12)	<0.001
Age at diagnosis, y	1.13	(1.13, 1.13)	<0.001
PSA level, ng/mL	0.87	(0.86, 0.88)	<0.001
Gleason 7 vs 2–6	0.85	(0.82, 0.89)	<0.001
Age × PSA	1.00	(1.00, 1.00)	<0.001
Age × (Gleason 7 vs 2–6)	0.99	(0.99, 1.00)	<0.001
PSA × (Gleason 7 vs 2–6)	1.01	(1.00, 1.02)	0.05

[*] Covariates were scaled so the intercept corresponds to a patient age 50 y with PSA 4 ng/mL and Gleason 2–6 at screen detection.